THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository: http://go.warwick.ac.uk/wrap

A Thesis Submitted for the Degree of PhD at the University of Warwick

http://go.warwick.ac.uk/wrap/57320

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

Transcription Factor Interactions at the Promoter of the *Arabidopsis* Circadian Clock Gene *LHY*

Siân Elizabeth Wynne Davies

A thesis submitted to the Department of Life Sciences in fulfilment of the requirements for the degree of Doctor of Philosophy

January 2013

CONTENTS

PAGE

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

LIST OF ABBREVIATIONS

ABA: Abscisic Acid

ABF: ABA Responsive Elements Binding Factor

ABI: ABA Insensitive

ABRE: Abscisic Acid Responsive Element

AD: Activation Domain

ANT: Aintegumenta

APB: Active Phytochrome Binding

APRR: *Arabidopsis* Pseudo-Response Regulator

ARF: Auxin Response Factor

BLAST: Basic Local Alignment Search Tool

BPC: Basic Pentacysteine Protein

BR: Brassinosteroid

bZIP: Basic Region/Leucine Zipper Motif

CAB: Chlorophyll A/B Binding Protein

CAT3: Catalase 3

CBF: C-repeat/DRE-Binding Factor

CBS: CCA1 Binding Site

CCA: Circadian Clock Associated

CCD: Charge-Coupled Device

CCR2: Cold, Circadian Rhythm and RNA Binding 2

cDNA: Complementary DNA

CDS: Coding DNA Sequence

CHE: CCA1 Hiking Expedition

ChIP: Chromatin Immunoprecipitation

Col-0: Columbia ecotype of *Arabidopsis thaliana*

D_{ex}: Dexamethasone

DPBF: Dc3 Promoter-Binding Factor

EE: Evening Element

EEL: Enhanced EM Level

ELF: Early Flowering gene

EMSA: Electrophoretic Mobility Shift Assay

ERF: Ethylene Responsive Factor

FKF: Flavin-Binding Kelch Repeat F-box

FLC: Flowering Locus C

FT: Flowering Locus T

GA: Gibberellin (Gibberellic Acid)

GARP: Gibberellin Response Protein

GBF: G-Box Binding Factor

GFP: Green Fluorescent Protein

GI: Gigantea

IAA: Indole-3-Acetic Acid Response Element

JA: Jasmonate (Jasmonic Acid)

LB: Lysogeny Broth media

LD: Light/Dark daily cycles

Ler-0: Landsberg erecta ecotype of *Arabidopsis thaliana*

LHY: Late Elongated Hypocotyl

LTAH: Leucine, Tryptophan, Alanine, Histidine

LUX: Lux Arrhythmo (aka PCL1)

MADS: MCM1, Agamous, Deficiens and SRF

ME: Morning Element

MeJA: Methyl Jasmonate

mRNA: Messenger RNA

MS0: Murashige and Skoog media

mY1H: Modified Yeast One-Hybrid Assay

NAM: No Apical Meristem Protein

OD_{600}: Optical Density/absorbance at a wavelength of 600nm

PCL1: Phytoclock 1 (aka LUX)

Pfr: Phytochrome (Far-red light absorbing form)

PHYA: Phytochrome A

PHYB: Phytochrome B

PIF: Phytochrome Interacting Factor

PIL: Phytochrome Interacting Factor3-Like

Pr: Phytochrome (Red-light absorbing form)

RNAi: RNA Interference

ROS: Reactive Oxygen Species

SA: Salicylic Acid

SAR: Systemic Acquired Resistance

SD: Synthetic Defined basic growth media for yeast

SD-LTAH: SD media lacking leucine, tryptophan, adenine and histidine.

SOC1: Suppressor of Over-Expression of CO1

SORLIP: Sequences Over-Represented in Light-Induced Promoters

St Cl II: Strong Class II mutation in G-box flanking bases (ACCACGTGTC to GTCACGTGAC)

STK: Seedstick, an ovule developmental gene

T1ME: TOC1 Morning Element

TBS: TCP-Binding Site

TCP: Teosinte branched 1, Cycloidea and PCF

TF: Transcription Factor

TMG: TOC1 MiniGene

TOC1: Timing of CAB Expression 1

VIP: VirE2-Interacting Protein

VirE2: *Agrobacterium* protein involved in DNA transfer during *Arabidopsis* infection

Wk Cl II: Weak Class II mutation in G-box flanking bases (ACCACGTGTC to CTCACGTGAG)

WS: Wassilewskija ecotype of *Arabidopsis thaliana*

Y1H: Yeast One-Hybrid

ZT: Zeitgeber Time, time since dawn

ZTL: Zeitlupe

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Isabelle Carré, and past and present members of the lab group, with a special thanks to Sally Adams, for all their help and encouragement. I must also give a big thank you to Peijun Zhang and Claire Hill for their generous help with the yeast assays, and for making me feel so welcome over at HRI. I'd also like to thank Hannah Lee, Holly Baum and the rest of the old C30/46 crowd for all the tea, gossip and cake based light relief over the past 4 years! Finally, I have to give extra special thanks to Christopher Howell for his endless help and support over the past two years.

DECLARATION

I declare that the work presented in this thesis was conducted by me under the direct supervision of Doctor Isabelle Carré, with the exception of those instances where the contribution of others has been specifically acknowledged. None of the work presented here has previously been submitted for any other degree.

Siân Elizabeth Wynne Davies

ABSTRACT

The circadian clock is the endogenous mechanism by which a wide variety of biological processes are regulated in anticipation of daily changes in the external environment. In *Arabidopsis thaliana*, the clock comprises a number of complex gene and protein interactions, involving multiple regulatory feedback loops. The clock gene *LHY* has a central role in these loops, activating and repressing morning- and evening-expressed genes respectively. These clock genes in turn sequentially repress the expression of *LHY* throughout the day and night, restricting it to a sharp transcriptional peak at dawn. However, the molecular mechanisms of these regulatory interactions with the *LHY* promoter were not known. Therefore, this project first aimed to determine which promoter motifs are responsible for mediating regulation of *LHY* circadian expression. This was achieved through *luciferase* assays with mutated *pLHY:LUC* reporter constructs, which identified the CT-rich region as responsible for rhythmic expression of *LHY*, and the G-box as mediating regulation by the clock protein TOC1. Since few regulators were known to target the *LHY* promoter, this project also aimed to identify transcription factors binding the promoter using a Yeast One-Hybrid assay. Transcription factors with roles in a wide variety of biological pathways were identified from this screen, with abiotic stress and plant defence pathways particularly well-represented. In addition, a number of antagonistic and synergistic regulatory interactions were established as occurring between stress factors and clock proteins at specific promoter motifs. We can therefore conclude that *LHY* is regulated by a complex network of transcription factor interactions, enabling the rapid integration of environmental stress signals into the clock.

CHAPTER 1

Introduction

1.1 – The Circadian Clock

Circadian Clocks are found in the vast majority of organisms; they enable an organism to prime its internal biological processes in advance of predictable daily and seasonal changes in the external environment. For example, the timing of the daily sleep-wake cycle in animals, and preparation for photosynthesis at dawn in plants are all regulated by their respective circadian clocks (Huang et al., 2011, Sehgal et al., 2007, McClung et al., 2000). The circadian clock itself is the central oscillator driving these daily biological rhythms and will be re-set or 'entrained' every day by external cues or 'zeitgebers' such as the onset of dawn. The term 'circadian' refers to the daily rhythms of these biological processes, which were discovered to cycle with a period of approximately 24 hours, equivalent to one day even when under constant conditions lacking any potential entrainment cues. The daily entrainment of the clock by environmental signals is necessary to allow it to synchronise to the 24 hour cycle and to adjust to changing conditions, e.g. the timing of dawn and day-length, across the year and so remain synchronised with the organism's environment.

The clock influences a wide range of biological processes in eukaryotes, from early development to energy metabolism, growth, immunity and reproduction (Panda et al., 2002, Covington & Harmer, 2007, Huang et al., 2011, Elliott et al., 1972). The number of genes regulated by the clock is correspondingly large, with around 10% of mammalian genes and at least 30% of plant genes believed to be directly controlled by the clock (Panda et al., 2002, Covington et al., 2008). In addition, circadian clocks

have been found in almost all organisms tested, with various species of archaea, prokaryotes, invertebrates, and fungi also displaying circadian rhythms (Bell-Pedersen et al., 2005). The only organisms currently known to lack circadian clocks are those living in environments where regular daily rhythms of light or temperature are not found, such as within the Arctic Circle (Lu et al., 2010). Circadian clocks therefore appear to confer an evolutionary advantage on organisms living in a world of cycling environmental conditions. In support of this, perturbation of the circadian clock has been linked to diseases related to energy metabolism in mammals. Correspondingly, the effectiveness of related disease treatments has been found to improve when administered at specific times of day (Li et al., 2012b). Understanding regulation of the clock in mammals and how this affects downstream processes is therefore a major field of study.

However, the study of circadian clocks in plants is no less important. Plants are largely immobile and are therefore far more vulnerable to changes in their local environment; they cannot choose where they grow so must be able to both accurately anticipate and prepare for seasonal changes, and also adapt these processes to the local environment in which they find themselves, which may be shaded, water-logged or drought-prone. Consequently, while the plant circadian clock directly controls the expression of a large percentage of the genome, it must also itself be controlled by a wide array of input signals. This greater control allows the clock to play a role in regulating the timing of such diverse plant processes as cell elongation, growth, flowering, germination, chlorophyll production, leaf movements and stomatal opening (McClung, 2006).

Understanding how these environmental signals feed into and regulate the plant clock and how the clock then translates these signals to effect complex changes in multiple downstream biological processes, is therefore an important area of study. This is particularly true for our understanding of processes such as the timing of flowering and growth, with their implications for crop yields.

1.1.1 – Overview of the *Arabidopsis* Circadian Clock

The circadian clock in *Arabidopsis thaliana* (a model organism from the agriculturally relevant Brassica family of plants) has been studied extensively over the past two decades. The *Arabidopsis* clock is composed of a number of complex gene and protein interactions, involving multiple feedback loops. The network of interactions between the clock oscillator genes alters throughout the day depending on which of these proteins are present and therefore which feedback loops are active. The rhythmic expression of these oscillator genes then regulates the expression patterns of a variety of clock-controlled genes, so generating 'output' signals. These can control the timing of metabolic and biological processes such as chlorophyll production, leaf movements, growth and flowering (McClung, 2006).

1.1.2 - LHY and CCA1

LHY (*Late Elongated Hypocotyl*) and CCA1 (*Circadian Clock Associated-1*) are two closely related MYB-family transcription factors. They have highly similar diurnal (in light/dark cycles) and circadian (in constant light) expression profiles, with their expression peaking around one to two hours after dawn. LHY and CCA1 were initially identified as important to the function of the *Arabidopsis* circadian clock through a series of gene knockout and overexpressor studies, detailed below (Schaffer

et al., 1998, Wang and Tobin, 1998, Green and Tobin, 1999). A central role for both LHY and CCA1 in the *Arabidopsis* circadian clock was later confirmed with a double *lhy/cca1* loss of function mutant (Mizouchi et al., 2002, Alabadi et al., 2002).

LHY was first identified as a clock-regulated gene whose constitutive overexpression abolished both rhythmic expression of other clock-regulated genes, and also its own endogenous expression pattern. It was therefore suggested as a central component of the circadian clock. Over-expression of *LHY* was also found to cause late flowering of plants, independent of photoperiod, elongated hypocotyls and reduced chlorophyll content (Schaffer et al., 1998).

CCA1 was first identified as binding the promoter of the light-harvesting chlorophyll a/b gene *Lhcb1*3* (*CAB1*), through the targeting of the binding sequence AAAAATCT (CCA1 binding site, CBS) by its N-terminal MYB domain. CCA1 expression was found to be transiently induced by light, since transcript levels increased on transfer to red light of dark-grown seedlings. In addition, CCA1 RNAi plants showed a reduced red light induction of *CAB1*, but not another phytochrome-induced gene. Therefore, CCA1 was suggested to have a specific role in mediating phytochrome-dependent light induction of *CAB1* (Wang et al., 1997). In addition, over-expression of *CCA1* was found to disrupt the circadian expression of the clock-regulated genes *CAB2*, *CCR2* and *CAT3*. *CCA1* over-expression also caused reduced and arrhythmic expression of both itself and *LHY*. Plants over-expressing *CCA1* had a similar phenotype to those over-expressing *LHY*, with long hypocotyls and late flowering. These findings suggested that both *LHY* and *CCA1* were central to the circadian clock (Wang & Tobin, 1998).

These similarities led to suggestions of functional redundancy between LHY and CCA1. This was supported by the observation that neither LHY nor CCA1 inactivation alone in mutant lines could completely abolish rhythmic expression of clock-controlled genes. However, it was also clear that *LHY* and *CCA1* were only partially functionally redundant, since neither was able to entirely compensate for the loss of the other, instead both producing short-period rhythms (Green & Tobin, 1999, Mizoguchi et al., 2002). It was therefore suggested that LHY and CCA1 may act as heterodimers as well as homodimers to regulate target genes, a proposal supported by the discovery that both proteins can bind to the clock-regulated *CAB2* and *TOC1* promoters (Green & Tobin, 1999, Alabadi et al., 2001). A shared binding site termed the Evening Element (AAAATATCT) was found to mediate the binding of both proteins to the *TOC1* promoter *in vitro*. This led to the first model for a regulatory negative feedback loop of the clock: with LHY and CCA1 found to repress *TOC1* expression, and TOC1 required for expression of *LHY* and *CCA1* (Alabadi et al., 2001). As described in Section 1.1.3.3, this initial model was later revised to include direct repression of *LHY* and *CCA1* expression by TOC1.

Key regulatory roles for LHY and CCA1 as part of the circadian clock were confirmed in *lhy/cca1* double loss of function mutants, which caused a loss of rhythmic expression of clock-regulated genes within two days of constant light conditions (Mizoguchi et al., 2002, Alabadi et al., 2002). However, the existence of other oscillator components was also predicted from these results, since advanced phase/short-period rhythms were still seen in the double mutants under diurnal and initial circadian conditions respectively.

1.1.3 - The PRR Transcription Factors

A number of transcription factors were identified concurrently with LHY and CCA1 as having a role in the circadian clock. These included the APRR (*Arabidopsis* PSEUDO-RESPONSE REGULATOR) transcription factors: APRR9, APRR7, APPR5, APRR3 and APRR1 (TOC1).

1.1.3.1 - TOC1 and the PRRs

TOC1 (*TIMING OF CAB EXPRESSION 1*) was originally identified by Millar et al. (1995) from a screen for period-length mutants in *Arabidopsis* using a *cab2::luciferase* reporter plant line. The *toc1-1* mutant plant line was found to shorten the circadian period of both reporter gene and primary leaf movements by 2-3 hours (Millar et al., 1995). Because this period shortening was shown to occur independently of light input and because *toc1-1*'s effects were seen across many different developmental pathways, TOC1 was proposed as part of the circadian oscillator (Somers et al., 1998).

Once its sequence was cloned and verified (Strayer et al., 2000), it became apparent that *TOC1* was the same gene as *PRR1*, part of a family of genes whose expression was shown to be circadian and staggered across the day in the following sequence: *PRR9, 7, 5, 3* and *TOC1/PRR1* (Matsushika et al., 2000). These genes were found to be expressed rhythmically in the same sequence and at the same relative intervals irrespective of day-length. Under 12L:12D conditions, mRNA levels of the PRRs peaked as follows: PRR9 at ZT-2, PRR7 at ZT-6, PRR5 at ZT-8, PRR3 at ZT-10 and TOC1 at ~ZT-12. Although PRRs 9, 7, 5, and 3 had sharp transcriptional peaks, the peak of TOC1 at ZT-12 was much broader, with mRNA levels high for several hours,

from ZT-10 to ZT-18 under 12L:12D conditions (Matsushika et al., 2000). Protein levels of TOC1 peaked later, at around ZT-18, but still displayed a broad peak throughout the evening, though levels rapidly decreased by ZT-22. This pre-dawn destabilisation of TOC1 protein was found to be dependent on the ZEITLUPE (ZTL) protein, which binds TOC1 protein and targets it for proteasomal degradation in the late night (Más et al., 2003b).

The first of the PRRs to be expressed after dawn is PRR9. This was also the only PRR gene to be rapidly and transiently induced by white light or red light pulses (Makino et al., 2001). Further experiments with red and far-red light led to the suggestion that PRR9 might be a bridging factor between light input signals from phytochromes and the central oscillator (Makino et al., 2001). Expression of PRR9 was also suggested to be repressed by TOC1, since mRNA levels of PRR9 were reduced to almost undetectable levels in plants constitutively over-expressing TOC1 (Makino et al., 2002). Again PRR9 appeared unique amongst the PRRs in this respect, since overexpression of TOC1 had only a minor negative effect on mRNA levels of PRRs 7, 5, and 3, which retained their circadian rhythms of expression.

1.1.3.2 - Roles of PRR9, 7 and 5 within the Clock

Through studies using a triple *prr9/7/5* knockout mutant, PRR 9, 7 and 5 were suggested to act semi-redundantly close to the central oscillator, down-regulating *LHY* and *CCA1* expression (Nakamichi et al., 2005). In turn, LHY and CCA1 were shown to activate the expression of *PRR9* and *PRR7* through direct binding to their promoters (Farré et al., 2005). Further studies found that PRR9, 7, and 5 were directly repressing *LHY* and *CCA1* during the day, through binding of their promoters around

the G-box (Nakamichi et al., 2010). PRR9 and PRR7 were therefore proposed to act in a negative feedback loop with LHY and CCA1.

PRR9, 7 and 5 were also found to up-regulate *TOC1* expression (Nakamichi et al., 2005). TOC1 was found to down-regulate the expression of all the PRRs, including itself, to varying degrees. However, the strongest repression by TOC1 was seen on *PRR9* (Makino et al., 2002, Strayer et al., 2000). PRR9, 7 and 5 were therefore proposed to be involved in a second negative feedback loop, with TOC1.

1.1.3.3 - Role of TOC1 within the Clock

Unlike the other PRR-family proteins, TOC1 was originally thought to be activating *LHY* and *CCA1* expression. In addition, LHY and CCA1 were found to negatively regulate *TOC1* through direct binding to its promoter. TOC1 was therefore proposed to act in a regulatory negative feedback loop with LHY and CCA1 (Alabadi et al., 2001).

Since TOC1 was not at that time known to possess a DNA-binding domain, it was presumed to act through an intermediate transcription factor, termed 'X'. One such candidate was PIF3, a transcription factor in the phytochrome signalling pathway that was found to be capable of binding the promoters of *LHY* and *CCA1*, and also to bind TOC1 protein *in vitro* and in yeast (Makino et al., 2002). Another potential intermediary was identified as *CHE* (*CCA1 HIKING EXPEDITION*), a TCP transcription factor also known to be capable of interacting with TOC1 protein. CHE was found to repress *CCA1* expression *in planta* through direct interaction with the TCP-binding site (TBS, GGNCCCAC) of the *CCA1* promoter. A regulatory feedback

loop was established when CCA1 was found to negatively regulate *CHE* expression and to bind the *CHE* promoter through its CBS motif. In addition, ChIP data showed TOC1 associating with the *CCA1* promoter around the TBS. As TOC1 was not thought able to bind DNA itself, it was therefore suggested that TOC1 interacted with CHE at the TBS to antagonise CHE's repression of *CCA1*. However, CHE was not found to bind the *LHY* promoter, suggesting that there was a different mechanism for the regulation of *LHY* by TOC1 (Pruneda-Paz et al., 2009).

It was subsequently discovered that TOC1 was capable of binding DNA directly through its CCT domain, which is conserved across the PRRs (Gendron et al., 2012). Purified TOC1 protein was found to bind to TGTG (T1ME) sequences from the *LHY* and *CCA1* promoters *in vitro*, and the CCT domain was found to be sufficient for binding. Notably, this T1ME sequence forms part of the Morning Element (ME, GTGTGG), which is enriched in the promoters of morning-expressed genes (Michael et al., 2008). ChIP-Seq on TOC1 Minigene (TMG) plants (*TOC1p::TOC1:YFP* in a *toc1-2* mutant) revealed many *in planta* binding targets of TOC1, including oscillator genes such as *LHY*, *CCA1*, *PRR9*, *PRR7*, *ELF4* and *LUX* (Huang et al., 2012). Binding of TOC1 to these promoters was rhythmic, peaking at ZT-15. Analysis of this data identified G-box-expanded ((a/c)C(a/t/g)CG(t/c)) and EE-expanded ((a/t/g)AA(t/g)ATC(t/g/c)) motifs as being enriched in TOC1-bound oscillator genes (Huang et al., 2012). In addition, microarray analysis of alcohol-induced TOC1 plant lines also identified sequences that were enriched in genes that were up-regulated by TOC1, including the G-box and GA-repeat element, or down-regulated by TOC1, including the TBS (Gendron et al., 2012).

Concurrent with these findings, TOC1 was found to be directly repressing *LHY* and *CCA1* expression. Alcohol-induction of TOC1 at various time-points under diurnal and circadian conditions resulted in reduced expression of *LHY* and *CCA1* (Gendron et al., 2012). Similarly, induced nuclear localisation of TOC1-GR through Dexamethasone (Dex) treatment of *35S:TOC1:GR* transgenic plants, caused reduced levels of *CCA1* transcript and *CCA1:LUC* activity (Huang et al., 2012). In addition, repression of *CCA1* was found to be dependent on TOC1 having a functional CCT domain (Gendron et al., 2012).

The role of TOC1 was therefore revised to be a DNA-binding negative component of the circadian oscillator (Huang et al., 2012). However, this re-classification of TOC1 as a direct repressor of *LHY* and *CCA1* appeared in conflict with some previous findings. Genetic data had indicated that TOC1 was activating *LHY* expression (Alabadi et al., 2001), and was required for reactivation of *CCA1* and *LHY* at dawn, since red-light induction of *LHY* and *CCA1* mRNA was deficient in *TOC1 RNAi* and *toc1-2* plants (Más et al., 2003a). However, in keeping with a repressive role for TOC1, constitutive overexpression of TOC1 had been shown to cause reduced levels and amplitude of circadian expression of *CCA1* and *LHY* mRNA (Makino et al., 2002, Huang et al., 2012).

The revised clock model (Huang et al., 2012), provided a possible explanation for the conflicting data on TOC1. The core loop of the clock in this model is composed of three inhibitory steps: (1) repression of *LHY* and *CCA1* by TOC1 and the PRRs, (2) repression of *TOC1* and the Evening Complex (EC) genes (Section 1.1.4) by LHY and CCA1, and (3) repression of *TOC1* and *PRR9* by the Evening Complex. This

style of oscillator driven by a triple-negative feedback loop, named the repressilator, had been previously established as a capable synthetic oscillator in *E. coli* (Elowitz & Leibler, 2000). The repressilator model (Figure 1.1B) enables indirect activation by transcriptional repressors through their repression of other repressors. For example, TOC1 can repress the other *PRR* genes, which would otherwise repress *LHY* and *CCA1* expression. Removal of TOC1 would therefore result in higher levels of the PRRs around the peak of *LHY*/*CCA1* expression, and consequently inappropriate repression of *LHY* and *CCA1*. TOC1 would therefore appear to be required for activation of *LHY* and *CCA1* at dawn, as well as their expression over the course of several circadian cycles.

The data for TOC1 fits this double negative indirect activation hypothesis, whereby acute exposure to TOC1 shows it to repress *LHY* and *CCA1* expression, but genetic data from *luciferase* assays taken when the clock has been cycling for several days shows TOC1 activating *LHY* and *CCA1* expression (Gendron et al., 2012, Huang et al., 2012, Alabadi et al., 2001, Más et al., 2003a). The repression of *LHY/CCA1* repressors such as PRR9, 7 and 5 by TOC1 is therefore a plausible explanation for the apparent activating effects of TOC1 on *LHY* and *CCA1* expression.

In conclusion, the PRRs and TOC1 are now known to target the *LHY* and *CCA1* promoters for repression. This repression is likely to occur sequentially throughout the day and evening, in accordance with the sequential expression profiles of the PRRs. In addition, these transcription factors are involved in interlocking negative feedback loops, whereby TOC1, PRR9, 7 and 5 repress *LHY* and *CCA1*, TOC1 also represses

expression of the *PRR*s, expression of *TOC1* is repressed by LHY and CCA1, and LHY and CCA1 activate *PRR9* and *PRR7* expression.

1.1.4 - The Evening Complex

Three other proteins were identified as having a regulatory role in the circadian clock. These were LUX, ELF3 and ELF4 which together form the Evening Complex. This Evening Complex was found to negatively regulate *PRR9, 7, 5* and *TOC1* through direct binding of their promoters.

1.1.4.1 – LUX, ELF3 and ELF4

EARLY FLOWERING 3 (*ELF3*) was identified as a transcription factor involved in photomorphogenesis and the photoperiodic regulation of flowering (Zagotta et al., 1996). It was also proposed to be involved in the transduction of light signals to the clock (Zagotta et al., 1996). In accordance with this, ELF3 was found able to physically interact with PHYB (Liu et al., 2001). In addition, the disruption of circadian rhythms in ELF3 knockout and overexpressor plants had been found to be both light- and phase-dependent, such that *elf3* plants were hypersensitive to light pulse-induced re-setting of the clock during the night, and plants overexpressing ELF3 displayed reduced sensitivity compared to wild type plants (Covington et al., 2001). ELF3 was therefore suggested to be involved in the gating of light input to the clock by antagonising the effects of light signals on the clock at night.

Expression of ELF3 was found to be circadian. However, circadian expression of ELF3 is not dependent on LHY, since rhythms were not disrupted by overexpression of *LHY* (Hicks et al., 2001). In contrast, CCA1 was found to bind to the *ELF3*

promoter to repress its expression, and CCA1 mRNA levels were found to be elevated in ELF3 overexpressing plants (Lu et al., 2012). Therefore, a negative feedback loop was established between ELF3 and the clock. This feedback loop was proposed to occur indirectly, through the *CCA1*-repressor PRR9, since ELF3 had been found to repress *PRR9* expression through direct binding of its promoter (Dixon et al., 2011).

EARLY FLOWERING 4 (*ELF4*) was also found to have a role in the PHYB signal transduction pathway, as well as in the regulation of both the circadian clock and timing of flowering (Doyle et al., 2002, Khanna et al., 2003). Like ELF3, it was also implicated in the evening-specific gating of light input to the clock. In addition, *ELF4* expression patterns showed a phase advance in *toc1* and *cca1/lhy* knockout plants, and ELF4 was found to be required for rhythmic circadian expression of *CCA1*, *LHY* and *TOC1*. It was therefore proposed to be a regulatory component of the clock oscillator (McWatters et al., 2007).

LUX ARRHYTHMO/PHYTOCLOCK1 (*LUX/PCL1*) was first identified as a GARP family MYB-related transcription factor from *CAB2:LUC* and *GI:LUC* screens for clock mutants (Hazen et al., 2005, Onai & Ishiura, 2005). Like ELF3 and ELF4, loss of *LUX* function results in an early flowering phenotype, and disruption of multiple circadian rhythms. Data from *lux* null mutants suggested that LUX represses the expression of *TOC1*, and is required for activation of *LHY* and *CCA1* (Hazen et al., 2005). LUX was also found to repress PRR9 expression through direct binding of its promoter (Helfer et al., 2011). In addition, expression of *LUX* is circadian-regulated (Onai & Ishiura, 2005) and the presence of an Evening Element in the *LUX* promoter suggested that this regulation may involve targeting of the promoter by LHY and

CCA1. LUX was also found to bind its own promoter *in vivo* to repress its own expression (Helfer et al., 2011).

1.1.4.2 - Role of the Evening Complex within the Clock

LUX, *ELF3* and *ELF4* were found to have highly similar mutant phenotypes and circadian expression profiles, with transcription peaking around subjective dusk. They were therefore tested against each other for protein interactions in yeast and *in planta*.

ELF3 was found to interact with both ELF4 and the C-terminal half of LUX in a Yeast Two-Hybrid assay. No interaction was detected between ELF4 and LUX. However, Yeast Three-Hybrid suggested a role for ELF3 in simultaneously binding both proteins to create a functional protein complex. This complex was confirmed *in planta*, when it was shown that endogenous ELF3 and LUX could be co-immunoprecipitated with ELF4-HA. In addition, the essential bridging role of ELF3 within this protein complex was confirmed in co-immunoprecipitation experiments using an *elf3-1* knockout line. Thus the ELF3-ELF4-LUX Evening Complex (EC) was established (Nusinow et al., 2011).

LUX, ELF3 and ELF4 had all been shown to regulate expression of *LHY* and *CCA1*, and ELF3 and LUX were also able to repress *PRR9* expression through direct binding of its promoter (Section 1.1.4.1). This targeting of the *PRR9* promoter by ELF3 and LUX revealed a functional role for the Evening Complex in the regulation of the clock, since the presence of LUX was required to recruit ELF3 to the *PRR9* promoter (Chow et al., 2012). In addition, *LUX*, *ELF3* and *ELF4* were all circadian-regulated, with their expression altered by *lhy* and *cca1* mutants (Section 1.1.4.1). The Evening

Complex was therefore suggested to be involved in a negative feedback loop within the clock, and was incorporated into the revised clock model (Huang et al., 2012).

1.1.5 - Models of the *Arabidopsis* Circadian Clock

In the Locke (2006) model (Figure 1.1A), the clock consisted of three interlocking regulatory feedback loops: (1) LHY/CCA1 activates the expression of *PRR9/PRR7* around dawn, and PRR9 and PRR7 jointly repress expression of *LHY*/*CCA1* during the day. Light signals feed into this loop to activate expression of *LHY* and *PRR9*, in line with experimental data. (2) TOC1 represses the light-activated component 'Y' in the evening, thought to be GI, which can activate *TOC1* expression. (3) LHY/CCA1 represses expression of both *TOC1* and *Y* during the day, and TOC1 activates *LHY/CCA1* via the unknown intermediate 'X'.

The primary differences between this earlier three-loop model (Gould et al., 2006) and the repressilator model (Huang et al., 2012) of the clock are that Pokhilko (2012) revised the role of TOC1 to be directly repressing *LHY*/*CCA1*, and incorporated the Evening Complex as an additional feedback loop.

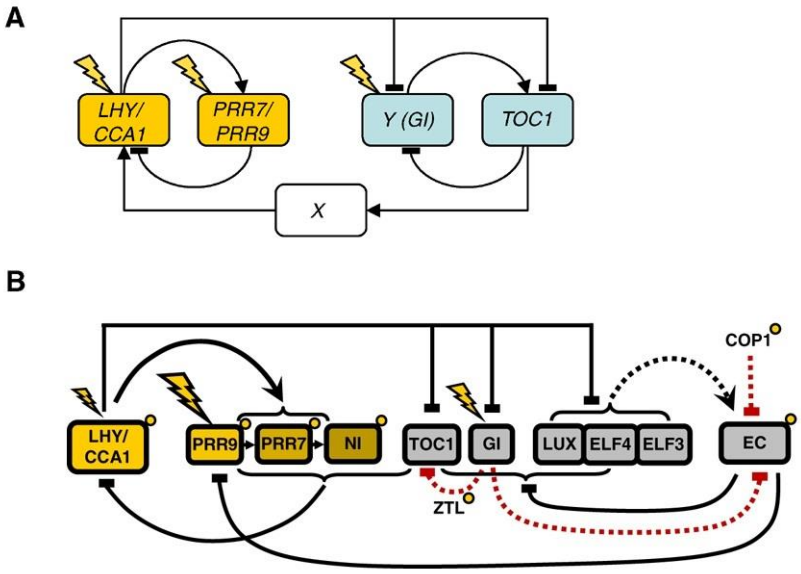


Figure 1.1: Models of the *Arabidopsis* circadian clock. A: The three-loop model (Locke et al, 2006). Regulatory interactions between genes (boxes) are indicated by arrows (positive effect) or blunt-ended (negative effect) symbols. Light input to genes is indicated by lightning symbols. Box colour indicates components of morning (yellow) or evening (blue) loops. B: The revised 'repressilator' model (Pokhilko et al, 2012), which has now superseded the three-loop model. Components of the morning and evening loops are shown in yellow and grey boxes respectively. Three protein components are included: ZTL, COP1 and the EC (Evening Complex). Transcriptional regulation is indicated by solid black lines, EC formation by a dashed black line, and post-translational regulation by dashed red lines (protein-mediated) or yellow circles (light-mediated). Acute transcriptional responses to light are indicated by lightning symbols.

The repressilator clock model (Figure 1.1B) left LHY/CCA1 in a negative feedback loop with PRR9, PRR7 and PRR5/NI, whereby LHY/CCA1 activates *PRR9/7/5* expression and the PRRs repress *LHY/CCA1* in the morning. The relationship between TOC1 and GI was revised in light of the discovery that GI mediates degradation of TOC1 through the stabilisation of ZTL protein (Kim et al., 2007). Both *TOC1* and *GI* are repressed by LHY/CCA1 during the day, and TOC1 completes this feedback loop by directly repressing expression of *LHY/CCA1* at night. The model places the Evening Complex interacting with both of these feedback loops. LHY/CCA1 represses *LUX*, *ELF3* and *ELF4* during the day. These form the Evening Complex (EC) at night which then represses two negative regulators of LHY/CCA1 in the form

of *TOC1* and *PRR9*. The EC is also antagonised by GI, and negatively regulates expression of its own component genes at night.

Therefore, the current clock model consists of three regulatory feedback loops: one in the morning (LHY/CCA1→ PRRs—| LHY/CCA1), one acting in the morning and night (LHY/CCA1—| TOC1—| LHY/CCA1) and another in the evening (LHY/CCA1—| EC—| TOC1 and PRR9).

1.2 - The *LHY* Promoter

At the time of the Locke (2006) model, little was known about regulation of the clock at a molecular level. However, it was clear from both mathematical models and experimental data (Section 1.1) that *LHY* and *CCA1* were central to the clock oscillator. Understanding the transcriptional regulation of these genes by other clock components was therefore fundamental to elucidating the molecular mechanisms by which the clock oscillator regulates itself. *LHY* and *CCA1* were also both known to be regulated by light, so were likely to be important in the integration of light signals into the clock and therefore in light-mediated entrainment of the circadian clock. The regulation of *LHY* and *CCA1* transcription was therefore considered an important area of study. The regulatory structure of the *LHY* promoter was examined by Spensley et al. (2009) (Section 1.2).

1.2.1 - Regulatory Regions of the *LHY* Promoter

Spensley et al. (2009) demonstrated through a 5' deletion analysis of the *LHY* promoter that the region starting at 957 basepairs upstream of the translational start

site (referred to as the -957/+1 *LHY* promoter) was sufficient for circadian and diurnal expression patterns of a *luciferase* reporter gene.

From this 5' deletion analysis, it was found that the *LHY* promoter region from 957 to 847 basepairs upstream of the translational start site (the -957/-847 promoter region) was essential for neither expression nor rhythmicity of the *LHY* promoter. Two conclusions could be drawn from this result: (1) that a positive regulator must act downstream of position -847 on the promoter in order to enable expression, and (2) that the -847/+1 promoter region (starting 847 basepairs upstream of the translational start site) was responsible for maintaining rhythmic expression.

The -957/-847 region of the *LHY* promoter was found to be responsible for regulating the phase of *LHY* expression in a photoperiodic-dependent manner. When entrained to either short days (8L:16D) or 12L:12D, the -847/+1 *LHY* promoter exhibited an advanced free-running phase of expression compared to the -957/+1 promoter. This suggested the presence of a repressor within the -957/-847 region acting to repress transcription in the late subjective night. However, this phase advance was not seen when plants were entrained to long days (16L:8D), indicating that the mechanism of the control of phase was more complex than a single repressor. It was therefore suggested that the -957/-847 region of the promoter might be targeted by waves of both transcriptional activators and repressors to regulate the phase of *LHY* expression.

The -957/+1 *LHY* promoter was therefore thought to consist of two broad functional regions, with the -957/-847 promoter region modifying the phase of *LHY* expression,

and the -847/+1 promoter region mediating both rhythmicity and activation of *LHY* expression.

1.2.2 - Functional Roles for Evolutionarily Conserved Promoter Motifs

Several evolutionarily conserved sequence motifs were identified within the -957/+1 *LHY* promoter using a comparative genomics technique (Picot et al., 2010), including a G-box, five novel AAAAA (5A) motifs, a CT-rich region and another three putative regulatory motifs (New Elements 1, 2 and 3) (Figure 1.2) (Spensley et al., 2009).

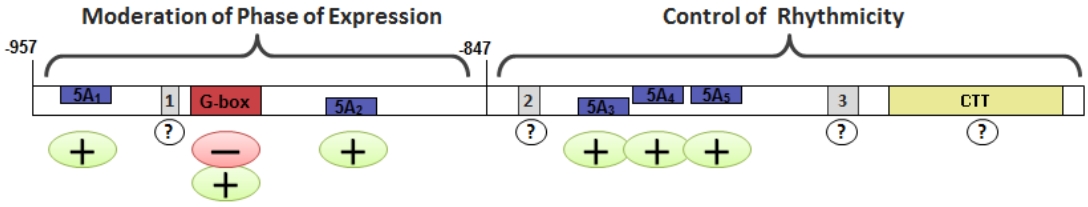


Figure 1.2: Regulatory regions and motifs of the *LHY* promoter. The distal promoter region (-957/-847bp upstream of translational start site) can modify the phase of LHY:LUC reporter expression, and the proximal promoter region (-847/+1) is sufficient for rhythmic expression of LHY:LUC reporter constructs. The 5A motifs are involved in activation of expression. The G-box may be targeted for both activation and repression. The function of Elements 1-3 and the CT-rich region is unknown.

Such conservation across species is highly suggestive of a functional role within the promoter, since for the sequence to be preserved there must be a strong functional imperative to protect against randomly accrued mutations within the motif sequence.

The individual contributions of these conserved promoter motifs to the mediation of promoter regulation were investigated by Spensley et al. (2009) using *luciferase* assays. Conserved motifs within the *LHY* promoter were individually mutated, and each mutated promoter cloned into a *luciferase* reporter vector. *In planta* expression of each LHY:LUC construct was then tracked and analysed over a period of days under differing environmental conditions. Two types of data were extracted from

these results: (1) mean expression levels, where absolute reporter gene expression level was averaged over several circadian cycles, and (2) expression patterns, where normalised expression levels were tracked over time to assess rhythmicity.

1.2.2.1 - The G-box

The G-box core hexamer (CACGTG) was already known to play a role in mediating responses to light (Martinez-Garcia et al., 2000), and its presence in this -957/-847 promoter region made it a prime target for further investigation. It was found by Spensley et al. (2009) using a series of point mutations in *LHY* reporter constructs that the amplitude of *LHY*'s expression was highly influenced by the flanking nucleotides around the G-box hexamer. The wider flanking sequences around the G-box were also shown to affect the binding of protein complexes to the promoter by electrophoretic mobility shift assays (EMSAs), suggesting that the promoter context of the G-box may be critical to its regulatory effects on the expression of *LHY*. The G-box was also implicated as a target for both activators and repressors of *LHY* expression, since mutations in the G-box flanking bases resulted in both a two-fold decrease in expression level and a reduction in the amplitude of oscillations, indicating an activating role for the G-box, and also a subtle broadening of the peak in constant light, indicating that the G-box moderates phase by repressing expression before and after the transcriptional peak at dawn.

1.2.2.2 - The 5A Motifs

Through a series of *luciferase* assays using mutated *LHY* promoter constructs, it was discovered that some or all of the 5A motifs were involved in the activation of *LHY* expression (Spensley et al., 2009). The five 5A motifs were found to be situated at

differing intervals along the promoter between positions -957 and -779 (Figure 1.2). They could be split into two distinct groups: the three proximal 5A motifs, clustered together in the -847/+1 promoter region between positions -779 and -805 (779bp and 805bp upstream of the translational start site), and the two distal 5A motifs which flank the G-box and Element 1 within the -957/-847 promoter region.

To uncover the role of the 5A motifs, reporter constructs were generated by Spensley et al. (2009), with a *luciferase* gene fused downstream of -957/+1 or -847/+1 promoters containing mutations in either the proximal or distal 5A motifs. *Luciferase* assays were performed in wild-type plants with the following constructs: -957 1,2m LHY:LUC, with the two distal 5A motifs ($5A_1$ and $5A_2$, collectively termed $5A_{12}$) disrupted in the -957/+1 promoter, and -957 345m LHY:LUC and -847 345m LHY:LUC, with the three proximal 5A motifs ($5A_3$, $5A_4$ and $5A_5$, collectively termed $5A_{345}$) mutated in the -957/+1 and -847/+1 *LHY* promoters respectively.

Luciferase assays using these constructs found that mutation of the proximal 5A motifs had no significant effect on expression of the -847/+1 promoter. However, mutations of either the distal or proximal 5A motifs in the context of the -957/+1 promoter were found to cause a two- to three-fold reduction in expression levels, a reduction in amplitude of expression and a broadening of the peak in constant light. The 5A motifs were therefore postulated to mediate activation of *LHY* expression. However, it remained unknown what the effect of losing all of the 5A motifs might be on the promoter's expression.

It was also implied from these results that the action of the 5A motifs may require the presence of one or more elements within the -957/-847 promoter region, since the effects on expression levels of mutating the proximal 5A motifs in the -957/+1 promoter were not seen when they were mutated in the -847/+1 promoter. The G-box was tentatively suggested as a possible focus of this interaction with the 5A motifs, since it had been shown through electrophoretic mobility shift assays (EMSAs) that protein binding at the G-box was negatively affected by unlabelled oligonucleotide competitors containing either of the two distal 5A motifs which flank the G-box. However, this was not conclusive evidence for a regulatory interaction between these motifs, so this proposed interaction requires further investigation.

1.3 - Aims

LHY is central to the function of the *Arabidopsis* circadian clock, playing a major and multi-faceted role in the negative feedback loops that drive rhythmic expression of the clock genes. The *LHY* promoter is targeted for regulation by other components of the clock oscillator, as well as being regulated by light. How this regulation is achieved at a molecular level, however, is largely unknown. The primary aim of the work presented in this thesis was therefore to elucidate the molecular mechanisms by which transcription factors regulate *LHY* expression as follows:

1) Identify functional roles for evolutionarily conserved *LHY* promoter motifs and investigate which of these motifs mediate activation and rhythmicity of *LHY* expression. (Chapter 3)

2) Identify *LHY* promoter motifs targeted by TOC1 for the regulation of *LHY* expression. (Chapter 4)

3) Identify transcription factors able to bind the *LHY* promoter, and map their binding to specific promoter motifs. (Chapter 5)

4) Investigate antagonistic and synergistic interactions between transcription factors binding at the *LHY* promoter. (Chapter 6)

CHAPTER 2

Materials and Methods

Unless stated otherwise, all standard laboratory techniques such as agarose gel electrophoresis, DNA ligation and ethanol precipitation of DNA were performed as described by Sambrook et al. (1989). Enzymatic manipulations of DNA were carried out as instructed by the enzyme suppliers. Sequencing of DNA was carried out by the University of Warwick Molecular Biology Service.

Extraction of DNA from agarose gels, purification of PCR reactions and extraction of plasmid DNA from *E. coli* were performed using the standard protocols of QIAgen Gel Extraction and PCR Purification Kits, and either QIAgen or Fermentas Miniprep Kits.

2.1 - Plant Lines

The TOC1 RNAi plants (in a WS background) used were as described by Más et al. (2003a) but with their *cab:luc* reporter constructs crossed out. The TOC1 overexpressor (TOC1ox WS) plant lines were obtained from Dr László Kozma-Bognár. The TOC1 MiniGene (TMG) plants (*pTOC1::TOC1:YFP* in *toc1-2* background) were as described by Más et al. (2003a). The following seed stocks were obtained from NASC (Nottingham Arabidopsis Stock Centre, (Scholl et al., 2000)): SALK_063665 (AT4G09180), SALK_021965 (AT2G41070), SALK_085497C (AT3G44460), SALK_144534C (AT4G36730), SALK_018426 (AT4G38900), SALK_003886 (AT2G21230), SALK_033320C (AT1G06850), SALK_140005C

(AT3G61180), N483900 (AT4G37750), N582906 (AT5G07110), SALK_078841 and SALK_016619C (AT3G12910), SALK_052716C (AT2G14210).

2.2 - Preparation and Transformation of Competent *E. coli* DH5α

Transformation of *E. coli* was performed using the recommended Invitrogen transformation protocol. Preparation of super competent *E. coli* was performed according to the protocol devised by the Chen Lab, Department of Chemical and Systems Biology, Stanford School of Medicine, described below.

Competent *E. coli* was prepared from a frozen stock of Invitrogen DH5α competent cells. 5ml of an LB overnight culture (grown at 37°C) was used to inoculate 500ml of SOB medium (2% w/v tryptone, 0.5% w/v yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$). The cell culture was grown in a 2L sterile flask at 37°C to an OD_{600} of 0.4-0.6, then chilled on ice for 10 minutes. All subsequent steps were performed at 4°C, with solutions and centrifuge tubes pre-chilled. The culture was centrifuged for 5 minutes at 4000 x *g*, and the pelleted cells were gently resuspended in 150ml of cold CMG buffer (50mM $CaCl_2$, 50mM $MgCl_2$). The cell suspension was incubated on ice for 15 minutes before the cells were pelleted again for 5 minutes at 4000 x *g*. The pelleted cells were gently resuspended in 36ml of cold CMG buffer and incubated on ice for 5 minutes. 1.26ml of high quality DMSO was added to the suspension, and mixed well. The suspension was incubated on ice for 5 minutes, then another 1.26ml of DMSO was added and the cell suspension incubated on ice for a further 5 minutes. The cell suspension (approx. 40ml) was gently

dispensed into aliquots in microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80°C.

2.3 - *Agrobacterium tumefaciens* Mediated Transformation of *Arabidopsis thaliana*

Competent *Agrobacterium* was prepared from an existing stock of a c58 pSoup containing strain. 2ml of an LB overnight culture was used to inoculate 50ml LB, which was grown at 28°C to an OD_{600} of 0.5-1.0. The culture was chilled on ice for 30 minutes then centrifuged at 3000 rpm, 4°C for 10 minutes. The cells were resuspended at 4°C in 1ml of 20mM $CaCl_2$ (pre-chilled), dispensed into 50µl or 100µl aliquots, frozen in liquid nitrogen and stored at -80°C.

Plasmid DNA was transformed into *Agrobacterium* as follows: 1µg DNA was added to 50-100µl frozen cells, thawed for 5 minutes at 37°C, frozen in liquid nitrogen and thawed again for 5 minutes at 37°C, and incubated with 1ml LB at 37°C for 2-4 hours. Cells were pelleted for 1 minute at 13,000 rpm, resuspended and plated on LB agar containing antibiotics (rifampicin 15 mg/l and appropriate selection for the plasmid). These plates were then incubated for 2-4 days at 28°C.

Transformation of plasmid DNA into *Arabidopsis thaliana* was performed using the floral-dip method as described in Zhang et al. (2006).

Transgenic plants were selected by either sowing on 50µg/ml Kanamycin MS0 plates, or by mixing the seed with sand and sowing on soil before spraying with 227µM Basta (glufosinate ammonium) 3 times per week for 2-3 weeks as appropriate.

2.4 - Surface Sterilisation of *Arabidopsis thaliana* Seed

For immediate use: Up to 150µl pre-dried seed in a 1.5ml eppendorf was soaked in 1ml 50% bleach/H20 containing 0.02% tween for 5 minutes. The seed was then washed 3 times in sterile water, and left to stand in the final wash until ready to plate on Murashige and Skoog (MS) agar (4.3g/l Murashige and Skoog basal salt mix (Sigma-Aldrich), pH 5.7, 1.5% (w/v) agar).

Short-term storage: As above with an additional final wash of 1ml 70-100% ethanol, with immediate and thorough drying on sterile filter paper before storage.

Vapour-phase sterilisation, for multiple lines: The method used was that of Clough and Bent (1998). 150µl of each seed line in open 1.5ml tubes was placed in a desiccator jar along with 100ml bleach. Immediately prior to sealing the jar, 3ml concentrated HCl was added to the bleach. Seeds were left sealed in fume hood for 5 hours.

2.5 - Luciferase Assays

Sterile *Arabidopsis* seed was sown on MS0 agar in 96 well plates, with approximately 15 seeds per well. Sterile water was added to the surrounding wells and the plate sealed with parafilm to prevent the media drying out. Plates were stored at 4°C for 4

days, then transferred to 22ºC incubator with a 12h:12h Light:Dark cycle for 7 days. Plants were sprayed with 5µM Luciferin (0.01% (v/v) Triton R X-100) on day 7 and returned to the incubator. Plates were moved to a photon-counting camera on day 8, where illumination was provided by custom-made red LED arrays. Luminescence was then monitored by digital imaging of plants using either the ORCAII c4742-98 CCD camera system (Hamamatsu Ltd., Welwyn Garden City, UK) or a liquid nitrogen cooled TEK 512x512DB CCD with an ST138 controller (Princeton Instruments Inc. Trenton, New Jersey). Automated imaging protocols, including control of illumination, were created and run using the MetaMorph™ software package (Molecular Devices Ltd, Wokingham, UK). Images were acquired with a 20 minute exposure every 2 hours for 6-7 consecutive days: 1 or 2 days under 12:12 LD conditions, and 4 or 5 days in constant light (specified individually for each experiment). Images were compiled and numerical data on luminescence intensity extracted from the 16-bit image files for analysis using MetaMorph v7.7.3.0. To correct for noise arising from the limitations of the cameras, background intensities were calculated for each image from regions containing no plants and subtracted from each data point from that image. This was repeated across every image in the time course. Temporal expression data for each transgenic line was normalised to its mean expression level in 72 hours of constant red light, and the mean and standard error of the mean taken at each time point across transgenic lines of the same construct/background type. Absolute expression levels in constant light were calculated from background corrected data, with each line corrected for number of seedlings. Differences between absolute mean expression values of *luciferase* constructs were assessed using two-tailed Student's T-tests. p-values of less than 0.05 were considered significant. Variance of samples was established using F-tests.

2.6 - RNA Extraction/cDNA Synthesis

Total RNA was extracted from seedlings using the Plant RNeasy kit (Qiagen) and contaminating genomic DNA removed by treatment with DNaseI (SIGMA). First-strand cDNA synthesis was carried out using Revert-aid H-Minus M-MuMLV Reverse transcriptase (Fermentas) and primed using random DNA hexamers.

2.7 - Chromatin Immunoprecipitation (ChIP)

Where stated in results chapters, seedlings were sprayed with a 25µM ABA solution to assess the effect of ABA on binding patterns of TOC1 to the *LHY* promoter. The method used for ChIP was adapted from Gendrel et al. (2002) as follows. For each chromatin preparation, approximately 500µl of seed was surface sterilised and sown on MS0 agar. Seedlings were grown for 2 weeks in 12L:12D at 22°C. Seedlings were harvested into a 50ml conical tube containing 20ml deionized water at the required timepoint; the water was drained through nylon mesh before cross-linking. Samples were cross-linked by vacuum infiltration in 20ml of 1% formaldehyde for 10 min, and the reaction stopped by adding 1.3ml of 2M glycine and vacuum infiltrating for a further 5 min. Samples were rinsed twice in deionized water, and frozen in liquid nitrogen for storage as necessary.

Each frozen sample was thawed on ice in 30ml of extraction buffer 1 (0.4M sucrose, 10mM Tris-HCl pH8, 5M beta-mercaptoethanol, 1mM PMSF and 1x protease inhibitors (Roche complete protease inhibitor cocktail)). Samples in buffer were ground over ice using a Polytron Status X120 (30 seconds, three times at 1 min intervals). Debris was removed by filtration through miracloth. Samples were

centrifuged at 4°C for 20 min at 1940 x g. The pellet was very gently resuspended on ice in 1ml of extraction buffer 2 (0.25M sucrose, 10mM Tris-HCl pH8, 10mM $MgCl_2$, 1% Triton X-100, 5M beta-mercaptoethanol, 1mM PMSF and 1x protease inhibitors), gently transferred to a 1.5ml tube and left on ice for 5-10 min. The nuclei were then collected by centrifugation at 4°C for 10 min at 14,000 x *g*. The nuclear pellet was resuspended on ice in 500µl of nuclei lysis buffer (50mM Tris-Hcl pH8, 10mM EDTA, 1% SDS, 1mM PMSF and 1x protease inhibitors). Samples were then sonicated on ice to give chromatin fragment sizes ranging from 100 to 1000bp (10 seconds at maximum amplitude, five times at 1 min intervals). Sonicated samples were centrifuged at 4°C for 10 min at 14,000 x *g*, the supernatant was transferred to a fresh tube and the sonication step was repeated. Samples were then centrifuged again (4°C for 10 min at 14,000 x *g*) and the supernatant transferred to a fresh 1.5ml tube. For each sample, 10% of the volume (~20µl) was removed as the Input DNA control, which did not undergo the following immunoprecipitation steps performed before the reversal of cross-linking.

For each sample, 125µl of chromatin was added to 1ml of ChIP dilution buffer (167mM NaCl, 16.7mM Tris-HCl pH8, 1.2mM EDTA, Triton X-100, 1mM PMSF and 1x protease inhibitors) on ice. Samples in buffer were pre-cleared by addition of 25µl protein A Dynabeads ® (Invitrogen) (beads were pre-equilibrated by rinsing three times in 1ml of ChIP dilution buffer, to a final dilution of 50%). Samples were incubated with beads on a rotating mixer wheel for 1 hour at 4°C, before application of a magnet to pellet the beads and transfer of supernatant to a fresh tube. Cleared samples were incubated overnight on rotating wheel at 4°C with either anti-LHY (1:200) or anti-GFP (1:1000) (abcam 290) antibodies. Samples were then incubated

with 50µl of equilibrated magnetic protein A Dynabeads ® (Invitrogen) on a rotating mixer wheel for 1 hour at 4°C, before application of a magnet. The pelleted beads were then washed to remove non-specific chromatin interactions (Haring et al., 2007) as follows. The beads were resuspended in 1ml low salt wash buffer (150mM NaCl, 0.2% SDS, 0.5% Triton X-100, 2mM EDT, 20mM Tris-HCl pH8), washed for 5 min at 4°C, and the supernatant discarded after pelleting with a magnet. Washing was repeated with 1ml of the following solutions for 5min at 4°C: four washes with high salt wash buffer (500mM NaCl, 0.2% SDS, 0.5% Triton X-100, 2mM EDT, 20mM Tris-HCl pH8), one wash with LiCl wash buffer (0.25M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1mM EDT, 10mM Tris-HCl pH8), and two washes with TE buffer (10mM Tris-HCl pH8, 1mM EDTA). Residual TE buffer was removed from samples after the final wash.

Before the reversal of cross-linking, 100µl of 10% Chelex resin (Biorad) was added to both the Input and the samples (Nelson et al., 2006). Input and samples were boiled for 10 minutes then cooled to room temperature, before digestion with 1µl of 20mg/ml Proteinase K for 30 minutes at 50°C. Input and samples were boiled for a further 10 minutes, then the Chelex resin was removed by centrifugation (5 min at 14,000 x *g*) at room temperature and the supernatant collected into a fresh tube. The pellet was washed in 100µl of TE buffer, centrifuged again, and this supernatant combined with the previous. Input and sample DNA was purified separately using mini-elute PCR columns (Qiagen). DNA was stored at 4°C.

2.8 - Quantitative PCR (qPCR)

To determine transcript levels of coding sequences from RNA extractions or enrichment of target promoters through ChIP, quantitative PCR was conducted using an ABI PRISM Sequence detection system (Applied Biosystems) and the Power SYBR® Green reagent (Applied Biosystems). Levels were calculated relative to the constitutively expressed gene *ACT2* (At3g18780). Each reaction was prepared in technical triplicate using 96 well plates and specific primer sets as required (Table 2.1). Differences between samples were assessed using paired Student's T-tests (for ABA treatment samples) or two-tailed T-tests (for transcript levels in knockout plants). p-values of less than 0.05 were considered significant. Variance of samples was established using F-tests.

Table 2.1: Primers for amplification of the *LHY* promoter.

Primer Name	Sequence
Primers along LHY promoter	
LHY -957 F LHY -985 R	5'CACTTTTACCTACGTGAGCTTC 5'TATCTCAAGTTGCTTCTCTACGA
LHY 5A-G F LHY 5A-G R	5'GAAGCAACTTGAGATATACCAAAAAGTG 5'GCAGATCGACACGTGGTGAT
LHY G-ele2 F LHY G-ele2 R	5'TCGATCTGCGATGACTTCTGTT 5'ATTGAAAAGTTTATTTGAGGCTGGAA
LHY ele2 F pLHY-2	5'CCAGCCTCAAATAAACTTTT 5' GACGGAATTCCCAGAAGCAATCTCAGC
Primers for amplifying -957/-754 region of LHY promoter	
pLHY-1 pLHY-2	5' CTGCGAATTCAAGCTTCTGGCTCGTAG 5' GACGGAATTCCCAGAAGCAATCTCAGC

2.9 - Site-Directed DNA Mutagenesis

Site-directed mutagenesis of promoter constructs was carried out using the QuikChangeTM system (Stratagene, La Jolla, CA, USA) with an additional 25 pre-PCR primer elongation step (Wang & Malcolm, 1999) to increase the efficiency of mutagenesis. Successfully mutated plasmids were identified by sequencing plasmid DNA extracted from transformed *E. coli* colonies.

Site-directed mutagenesis was performed to allow identification of binding sites of transcription factors on the *LHY* promoter in yeast, as well as for correction of sequence errors in transcription factor coding sequences (Tables 2.2 and 2.3).

Table 2.2: Primers for site-directed mutagenesis of the *LHY* promoter.

Motif Name	Motif Sequence	Mutated Sequence	Primer Name	Primer Sequence
Element 1	CAGCCACTA	CAGGTCCTA	LHY-ele1m F	5'-CAAAAAGTGCAGTAGACAGGTCCTACAATATCACCACGTG
			LHY-ele1m R	5'-CACGTGGTGATATTGTAGGACCTGTCTACTGCACTTTTTG
Element 2	AGCCTCAAATAAA	AGATCGAAATAAA	LHY-ele2 m1 F	5'-GGTGCTGTTCCAGATCGAAATAAACTTTTC
			LHY-ele2 m1 R	5'-GAAAAGTTTATTTCGATCTGGAACAGCACC
Element 2	AGCCTCAAATAAA	AGCCTCATCGCAA	LHY-ele2 m2 F	5'-CCTTGGTGCTGTTCCAGCCTCATCGCAACTTTTCAATTAAAATTTTTC
			LHY-ele2 m2 R	5'-GAAAAATTTTAATTGAAAAGTTGCGATGAGGCTGGAACAGCACCAAGG
Element 3	GTGGCTGAGATTGCTTC	GTGGCTCGATTTGCTTC	LHY-ele3 m1 F	5'-GGGAAAAATTGTTGTGGCTCGATTTGCTTCTGGGAATTCGAGC
			LHY-ele3 m1 R	5'-GCTCGAATTCCCAGAAGCAAATCGAGCCACAACAATTTTTCCC
Element 3	GTGGCTGAGATTGCTTC	GTGGCTGAGATTGAGGT	LHY-ele3 m2 F	5'-GTTGTGGCTGAGATTGAGGTTGGGAATTCGAGCTCAC
			LHY-ele3 m2 R	5'-GTGAGCTCGAATTCCCAACCTCAATCTCAGCCACAAC
G-box	ACCACGTGTC	ACCACCCGTC	G-box core m F	5'-GACAGCCACTACAATATCACCACCCGTCGATCTGCGATGACTTC
			G-box core m R	5'-GAAGTCATCGCAGATCGACGGGTGGTGATATTGTAGTGGCTGTC
G-box	ACCACGTGTC	GTCACGTGAC	St Cl II F	5'-GACAGCCACTACAATATCGTCACGTGACGATCTGCGATGACTTC
			St Cl II R	5'-GAAGTCATCGCAGATCGTCACGTGACGATATTGTAGTGGCTGTC
G-box	ACCACGTGTC	CTCACGTGAG	Wk Cl II F	5'-GACAGCCACTACAATATCCTCACGTGAGGATCTGCGATGACTTC
			Wk Cl II R	5'-GAAGTCATCGCAGATCCTCACGTGAGGATATTGTAGTGGCTGTC
5A motif ($5A_1$)	TACCAAAAAGT	TATGTCAAAGT	Site 1M F	5'-GCAACTTGAGATATATGTCAAAGTGCAGTAGACAGCCACTAC
			Site 1M R	5'-GTAGTGGCTGTCTACTGCACTTTGACATATATCTCAAGTTGC
5A motif ($5A_2$)	TGTTTTTTCCA	TGTTTTGACA	Site 2M F	5'-GATCTGCGATGACTTCTGTTTTGACAATTTATACCCTTGGTGTTCC
			Site 2M R	5'-GGAACACCAAGGGTATAAATTGTCAAAACAGAAGTCATCGCAGATC
5A motif ($5A_5$)	GGGGAAAAATT	GGTGTCAAATT	Site 345M1 F	5'-CCAAAAATTAGGTGTCAAATTGTTGTGGCTGAGATTGCTTCTGGC
			Site 345M1 R	5'-GCCAGAAGCAATCTCAGCCACAACAATTTGACACCTAATTTTTGG
5A motif ($5A_{3,4}$)	AATTTTTCCAAAAATT	AATTTACTGTCAAATT	Site 345M2 F	5'-GCCTCAAATAAACTTTTCAATTAAAATTTACTGTCAAATTAGGTG
			Site 345M2 R	5'-CACCTAATTTGACAGTAAATTTTAATTGAAAAGTTTATTTGAGGC

Table 2.3: Primers for re-amplification and correction of transcription factor coding sequences for yeast assays.

Primer Name	**Sequence**
Primers for site-directed mutagenesis	
ABI5 stop F	5' AAGAAAGCTGGGTGTCAGAGTGGACAACTC
ABI5 stop R	5' AGTTGTCCACTCTGACACCCAGCTTTCTTG
ABI3 stop F	5' CAAGAAAGCTGGGTGTCATTTAACAGTTTGAG
ABI3 stop R	5' CTCAAACTGTTAAATGACACCCAGCTTTCTTG
Primers for amplifying coding sequences from cDNA	
PIF 7 F	5’ GGCTTCACCATGGTCGAATTATGGAGTTAAAGAGCTC
PIF 7 R	5’ GCTGGGTGTCAATCTCTTTTCTCATGATTCGAAGAAC
PRR9 F	5’ GGCTTCACCATGGGGGAGATTGTGGTTTTAAGTAGTG
PRR9 R	5’ GCTGGGTGTCATGATTTTGTAGACGCGTCTGAATTCACG
PRR7 F	5’ GGCTTCACCATGAATGCTAATGAGGAGGGGGAGGGTT
PRR7 R	5’ GCTGGGTGTCAGCTATCCTCAATGTTTTTTATGTCGTTA
TOC1-N-F	5’ GGCTTCACCATGGATTTGAACGGTGAGTGTAAAGG
TOC1-N-R	5’ GGGTGTCAATTACTATTTCTTTTCATTGGCTCATG
TOC1-C-F	5’ GGCTTCACCATGAAAAGAAATAGTAATCCAGCGC
TOC1-C-R	5’ GGGTGTCAAGTTCCCAAAGCATCATCCTGAGGAG
Primers for introducing gateway compatible cloning sites	
attB1 F	5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATG
attB2 R	5’ GGGGACCACTTTGTACAAGAAAGCTGGGTGTCA
TOC1-1-F	5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGATTTGAACGGTGAGTGTAAAGG
TOC1-730-F	5’ GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGAAAAGAAATAGTAATCCAGCGC
TOC1-747-R	5’ GGGGACCACTTTGTACAAGAAAGCTGGGTGTCAATTACTATTTCTTTTCATTGGCTCATG
TOC1-1857-R	5’ GGGGACCACTTTGTACAAGAAAGCTGGGTGTCAAGTTCCCAAAGCATCATCCTGAGGAG

2.10 - Gateway Cloning

Where necessary, primers were designed to introduce flanking Gateway attB sites to target DNA using primer design and sequence specifications as defined by the Invitrogen Gateway Technology manual. Cloning was performed using standard Invitrogen Gateway Technology enzymes and vectors.

2.11 - Yeast Transformation

All yeast transformations were performed using the Lithium Acetate transformation method as described in Gietz & Schiestl (2007).

2.12 - Yeast One-Hybrid

Transcription factors able to bind the *LHY* promoter were identified with a Yeast One-Hybrid screen of a collection of 1181 *Arabidopsis* transcription factors (constructed from the REGIA clone library (Paz-Ares, 2002) by the PRESTA group, Warwick HRI). pDEST22 vectors containing coding sequences of individual transcription factors tagged with a yeast activation domain were transformed into AH109 haploid mating yeast in groups of 12 in 96 well plates. Two arrangements of the transcription factor library were tested simultaneously, each pooled into groups containing different combinations of transcription factors. Bait sequence was constructed by amplifying a fragment of the *LHY* promoter from 957 to 747 basepairs upstream of the translational start site. The promoter fragment was cloned into EcoR1 sites in a pHIS3LEU2 vector (containing Histidine reporter gene) and transformed into a haploid Y187 yeast strain. Libraries were plated (3µl of each pool) on double dropout selection medium (Synthetic Defined Medium -Leucine -Tryptophan, SD-LT) and allowed to dry at

room temperature before 3µl of the pLHY Y187 strain was plated on top of each library spot and allowed to dry. Plates were grown overnight at 28°C. Velvets were used to remove excess growth, and SD-LT plates were printed onto SD-LTH selection plates (*-leu-trp-his*) and grown for a further 4 days. To compensate for auto-activation of promoters in yeast, inhibitors of histidine biosynthesis (3AT) were added during preparation of media at varying concentrations between 0mM and 100mM. Growth was compared across selection plates and was indicative of transcription factor binding to the promoter.

2.13 - Modified Yeast One-Hybrid

For detailed methodology of the modified Yeast One-Hybrid assay, see Section 6.2.1. Diploid yeast cultures were sequentially transformed with two transcription factors in pDEST22 and pARC352 vectors and the -957/-754 region of the *LHY* promoter in a pHISLEU vector. Cultures were grown in 1ml liquid SD-LTA media in 96 deep-well plates at 30°C and 200rpm for 3 days. Cell concentrations were determined from these cultures using optical density (OD) measurements taken at 600nM and a pre-existing cell count to OD reference sheet (PRESTA group, University of Warwick). 10^8 cells were removed from each culture and diluted in dH20 to a concentration of 10^7 and 10^6 cells/3µl. Each culture was plated on SD-LTA and SD-LTAH media containing 3AT inhibitors at a range of concentrations as described in Section 6.2.1. Yeast were grown for 3 days at 30°C and digitally photographed. Images were de-saturated and contrast values adjusted using the GNU Image Manipulation Program.

Transcription factors were sequenced to check for errors in the coding sequences. Primers for re-amplification and correction of coding sequences are in Table 2.3.

ABF3, NAM, FLC and LHY did not contain any nucleotide errors. However, MADS44 was found to contain two non-synonymous substitutions: Alanine to Valine (A191V) and Tyrosine to Histidine (Y103H). Alanine and Valine have highly similar structure and properties (Levy et al., 2001), so this substitution was considered unlikely to affect protein structure. The replacement of Tyrosine with a Histidine would involve a change in amino acid properties, from hydrophobic to hydrophilic (Bioinformatics for Geneticists. Ed: Barnes, M.R. and Gray, I.C. pp. 302-303). However, these errors were left uncorrected as neither substitution fell within the DNA-binding domain or dimerization domain of MADS44, and were therefore unlikely to affect any interactions it might be involved in at the *LHY* promoter.

The PIF7 sequence was found to contain a 5' truncation of its CDS, corresponding to a known and relatively unstudied splice variant. In addition, PRR9, PRR7 and TOC1 were found to contain significant errors. PRR9, PRR7, TOC1 and full-length PIF7 were re-amplified from *Arabidopsis thaliana* cDNA to correct their sequence errors and allow re-testing for binding to the *LHY* promoter within the mY1H assay. Due to difficulties experienced in expressing full-length TOC1 in yeast, it was amplified and tested as the N-terminal half 1-675bp downstream of translational start site (TOC1(N)) and the C-terminal half 657-1785bp downstream of translational start site (TOC1(C)) as well as full-length CDS (TOC1(F)). ABI3 and ABI5 contained several non-synonymous nucleotide errors and were corrected by site-directed mutagenesis. Therefore, although the initial Y1H screen used the splice variant of PIF7 and the mutated clones of PRR9, PRR7, TOC1, ABI3 and ABI5, all subsequent assays (mutated promoter assays and mY1H) used corrected clones for these transcription factors.

CHAPTER 3

Regulatory Roles of Conserved Motifs within the *LHY* Promoter

3.1 - Introduction

As described in the Introduction (Figure 1.2), it was determined by Spensley et al. (2009) that *LHY* upstream sequences starting 957 basepairs upstream (position -957) of the translational start site (position +1) were sufficient to drive rhythmic expression of a *luciferase* reporter gene with a phase similar to that of the endogenous *LHY* transcript. Two broad functional regions were identified within this -957/+1 *LHY* promoter: the proximal promoter region (positions -847 to +1) which was sufficient for rhythmic expression of *LHY*, and the distal promoter region (positions -957 to -847) which was involved in modifying the phase of this rhythmic expression.

Evolutionarily conserved sequence motifs were identified within the -957/+1 region of the *LHY* promoter: a G-box, five 5A motifs, a CT-rich region and three novel motifs, which were designated Elements 1, 2 and 3. Spensley et al. (2009) assigned functional roles to the G-box and 5A motifs in the regulation of *LHY* expression. The G-box, which was previously shown to play a role in light regulation of etiolated plants (Martinez-Garcia et al., 2000), was suggested to have a dual role in transcriptional activation and repression, so moderating the phase and level of *LHY* expression.

The 5A motifs were found to be involved in the activation of expression, but whether all five 5A motifs acted redundantly to activate expression of *LHY* was not studied. Mutation of the distal 5A motifs affected expression levels of (-957/+1) LHY:LUC. However, mutation of the proximal 5A motifs had no effect on (-847/+1) LHY:LUC expression levels, suggesting either that the 5A motifs are not functional within the -847/+1 region of the promoter, or that their action requires the presence of distal sequences. An interaction between the proximal 5A motifs and the distal promoter region was suggested by the authors, since although no effect of mutating the proximal 5A motifs was seen with truncated (-847/+1) LHY:LUC promoter constructs, they did affect the expression levels of full-length (-957/+1) LHY:LUC promoter constructs. Here we test the functionality of the proximal 5A motifs by comparing the effect of mutating the distal 5A motifs with that of mutating all 5A motifs.

In addition, although it was known that the proximal promoter region was sufficient for both rhythmicity and activation of *LHY* expression, it had not yet been identified which motifs mediated this regulation. As the proximal 5A motifs were not essential for expression, candidate motifs were the CT-rich region, Element 2 and Element 3, whose functionality was unknown.

The experimental aims were therefore as follows:

(i) Determine whether the proximal and distal 5A motifs act redundantly to regulate *LHY*'s expression (Section 3.2.1.1)

(ii) Clarify the nature of the interaction between the 5A motifs and distal promoter region (Section 3.2.1.2)

(iii) Characterise the role of the CT-rich region in the regulation of *LHY* expression (Section 3.2.2)

(iv) Investigate the functionality of evolutionarily conserved Elements 1, 2 and 3 (Section 3.2.3)

3.2 - Results

3.2.1 - The Role of the 5A Motifs in the Regulation of *LHY* Expression

3.2.1.1 - 5A Motifs are Required for the Expression of *LHY*

In order to test whether the 5A motifs function redundantly to regulate expression of *LHY*, a -957 12345m LHY:LUC construct (generated by Mark Spensley, Figure 3.1) with all five 5A motifs disrupted in the full-length promoter was transformed into wild-type (Wassilewskija, WS) plants to examine its *luciferase* expression pattern. Effects of mutating all 5A motifs were compared to the effects of mutating 5As in the distal region only. Multiple independent transgenic lines were grown for 7 days under cycling conditions of 12 hours white light and 12 hours of darkness, before spraying with a 5µM solution of Luciferin. The plants were then transferred to the CCD camera chamber and imaged every 2 hours for 2 days in 12L:12D in red light to examine diurnal rhythms, then for a further 5 days in constant red light to examine free-running rhythms. Expression levels for each transgenic line were calculated independently as the mean luminescence per seedling over 72 hours of constant light, beginning at the first subjective dawn. Background levels were determined at each time point and corrected for as described in methods (Chapter 2).

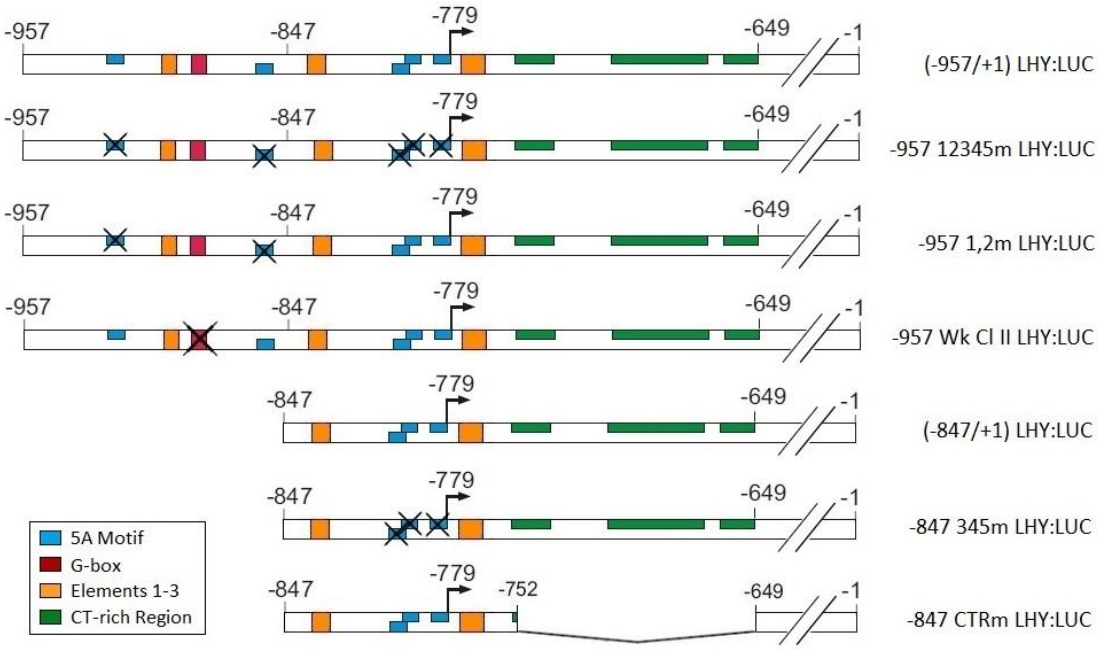


Figure 3.1: Mutated LHY:LUC reporter constructs. (-957/+1) LHY:LUC and (-847/+1) LHY:LUC reporter constructs were mutagenised by Mark Spensley to disrupt or remove individual *LHY* promoter motifs (Spensley et al., 2009).

The newly generated transgenic lines of -957 1,2m LHY:LUC did not show the significant reduction in expression compared to -(957/+1) LHY:LUC that was seen with previous transgenics (Spensley et al., 2009). This is most likely due to low sample size and high variability of -957 1,2m LHY:LUC. However, the simultaneous disruption of all 5A motifs in the -957 12345m LHY:LUC construct caused its expression to reduce significantly from that of the wild-type (-957/+1) LHY:LUC and -957 1,2m LHY:LUC constructs ($p<0.0001$ and $p<0.05$ respectively, Figure 3.2), to the extent of becoming too low for detection of circadian rhythms through *luciferase* assays. This result indicates that proximal 5A motifs are functional and act redundantly with distal motifs to mediate activation of *LHY* transcription.

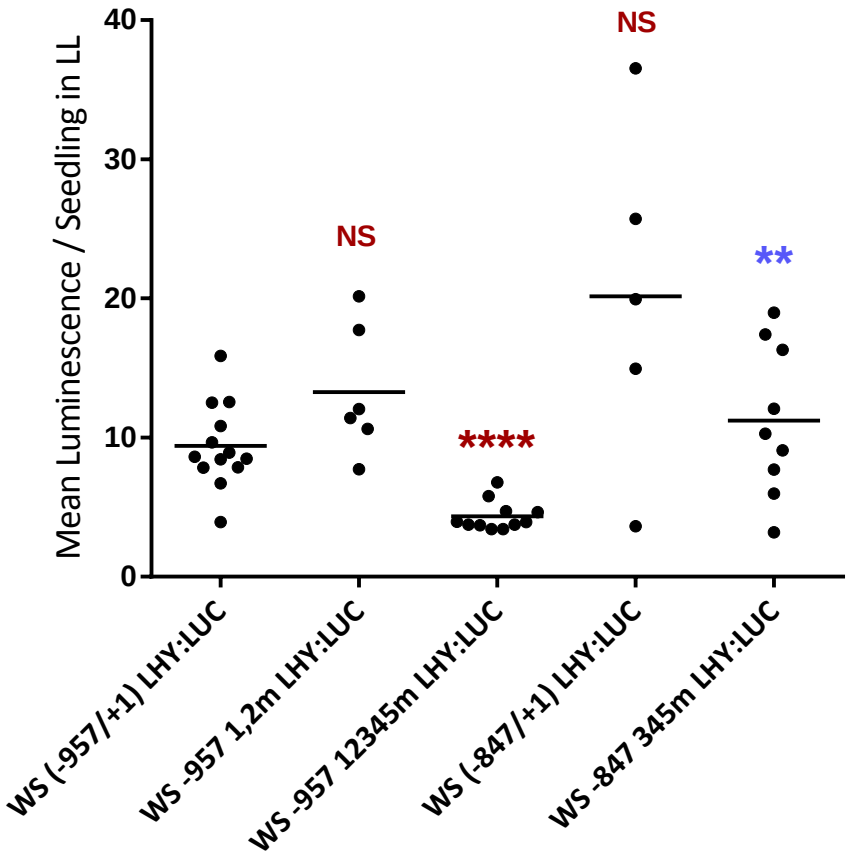


Figure 3.2: The 5A motifs are required for expression of full-length but not truncated LHY:LUC reporter constructs. Expression levels per seedling for each transgenic line were averaged over a period of 72 hours in constant red light (after 7 days entrainment to 12L:12D). Each data point represents the mean expression level from one transgenic line. Mean levels for each reporter construct are shown by horizontal lines. Red stars indicate p-values from T-tests comparing constructs to (-957/+1) LHY:LUC. Blue stars indicate a comparison between -957 12345m LHY:LUC and -847 345m LHY:LUC (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; NS = not significant).

3.2.1.2 - The Proximal 5A Motifs Interact with the Distal Promoter Region to Promote Expression

In a confirmation of results by Spensley et al. (2009), the mean expression level of -847 345m LHY:LUC was not found to be significantly different to that of either (-847/+1) LHY:LUC or (-957/+1) LHY:LUC (p>0.9 and p>0.05 respectively, Figure 3.2) when assayed as described in Section 3.2.1.1. Therefore, although the 5A motifs are essential for expression of the full-length promoter construct (Section 3.2.1.1) they

are not required for expression of the truncated promoter construct. This indicated a regulatory interaction between the proximal 5A motifs and the distal promoter region.

Furthermore, that deletion of the distal promoter region rescued the low expression levels caused by loss of the 5A motifs, with -847 345m LHY:LUC showing significantly higher levels of expression than -957 12345m LHY:LUC ($p<0.01$, Figure 3.2), suggesting that deletion of the distal promoter region removes a binding site for a transcriptional repressor.

However, as the deletion of the distal promoter region did not significantly increase expression of the wild-type construct (as would be expected from removal of an inhibitory sequence), we propose that this region also mediates the binding of an activator. In this hypothesis, deletion of the distal promoter region would result in removal of both activation and inhibition, and hence would not affect expression levels. Since the proximal 5A motifs are required for activation of transcription by the distal promoter region, the simplest hypothesis would be that these 5A motifs mediate binding of an activator. Another possibility is that proteins binding the 5A motifs act to prevent the action of a repressor binding the distal region of the promoter.

3.2.2 - Defining the Role of the CT-rich Region in the Regulation of *LHY* Expression

3.2.2.1 - The CT-rich Region Activates *LHY* Expression Redundantly with the Distal Promoter Region

The functionality of the CT-region was investigated by assaying reporter gene expression of full-length and truncated promoter constructs containing a deletion of

the CT-rich region: -957 CTRm LHY:LUC and -847 CTRm LHY:LUC (generated by Mark Spensley). Experiments were set up as described in Section 3.2.1.1.

The mean expression level per seedling of the -847 CTRm LHY:LUC construct was significantly lower than that of the (-847/+1) LHY:LUC promoter ($p<0.0001$, Figure 3.3), to the extent of becoming undetectable. This indicated that the CT-rich region is required for transcriptional activation through the -847/+1 promoter region.

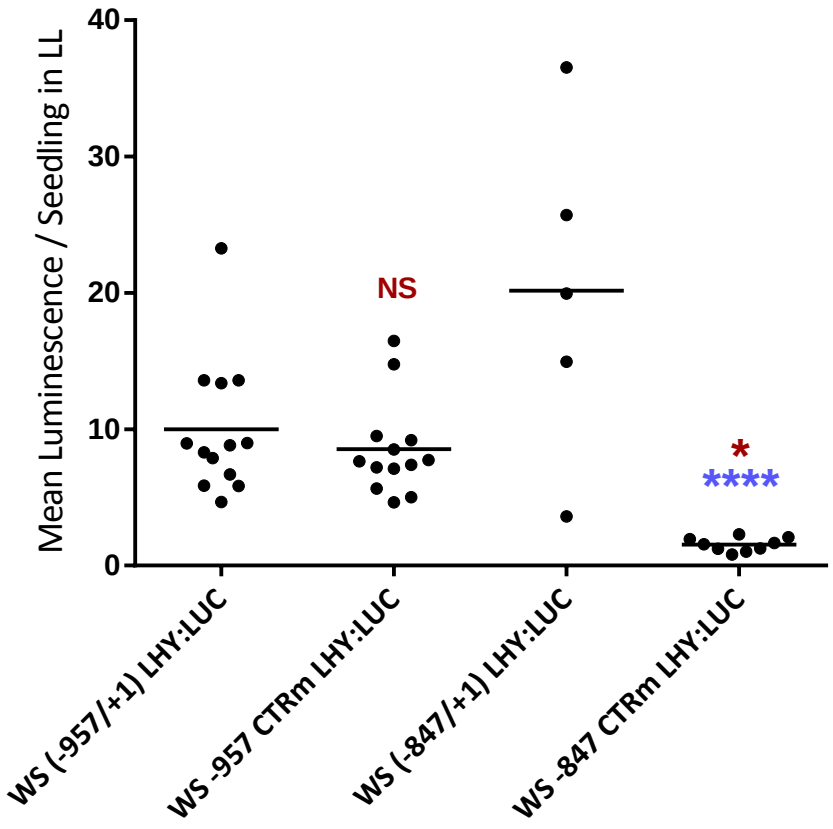


Figure 3.3: The distal region compensates for loss of CT-rich region activation. Expression levels per seedling for each transgenic line were averaged over a period of 72 hours in constant red light (after 7 days entrainment to 12L:12D). Each data point represents the mean expression level from one transgenic line. Mean levels for each reporter construct are shown by horizontal lines. p-values from T-tests are shown (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; NS = not significant). Red stars indicate comparison to wild-type construct: (-957/+1) LHY:LUC or (-847/+1) LHY:LUC. Blue stars indicate comparison between -957 CTRm LHY:LUC and -847 CTRm LHY:LUC.

However, mean *luciferase* expression level of the -957 CTRm LHY:LUC construct was not significantly different from the unaltered (-957/+1) LHY:LUC construct ($p=0.83$, Figure 3.3). Therefore, the CT-rich region is not required for expression when the distal promoter region is also present. These results indicate that (i) the CT-rich region mediates binding of transcriptional activators, and (ii) these transcriptional activators act redundantly with other activators binding the distal promoter region.

3.2.2.2 - The CT-rich Region Regulates Circadian and Diurnal Rhythmicity

None of the mutations or promoter deletions examined by Spensley et al. (2009) were able to abolish the rhythmic expression of *LHY*. However, the -847/+1 promoter region was shown to be sufficient for maintaining rhythmicity of expression.

To uncover which promoter elements within the -847/+1 region might be mediating rhythmicity, the temporal expression patterns of *luciferase* reporter constructs containing mutated elements within this region were investigated as described earlier. To enable direct comparison of waveforms between constructs, luminescence levels for each transgenic line were normalised to average levels in constant light.

When the expression pattern of -957 CTRm LHY:LUC was analysed, it was found that despite its unaltered expression level compared to the wild-type promoter construct (Figure 3.3), rhythmic expression of -957 CTRm LHY:LUC was abolished under both diurnal and free-running conditions (Figure 3.4). The CT-rich region is therefore essential for both circadian and diurnal rhythmicity of the (-957/+1) *LHY* promoter.

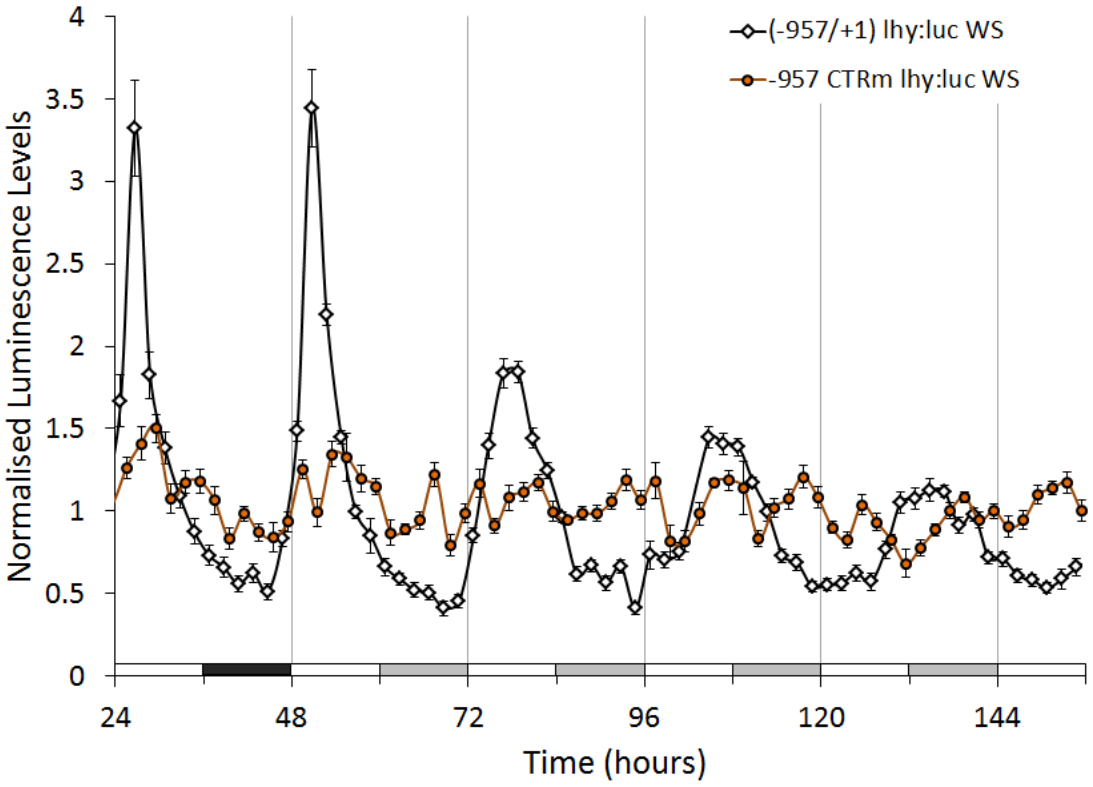


Figure 3.4: The CT-rich region is required for rhythmic expression. Plants were grown on MS0 agar for 7 days at 22°C in 12L:12D white light. They were then imaged for 2 days under 12L:12D of red light before transfer to constant red light. At least 6 independent transgenic lines were analysed per construct. Temporal patterns of luminescence were normalised to the mean expression level in constant light then averaged across independent transgenic lines. Error bars indicate the Standard Error of the Mean (SEM).

3.2.3 - Investigating Promoter Motif Interactions and the Functionality of Conserved Elements 1, 2 and 3

Whether the evolutionarily conserved motifs Elements 1, 2 and 3 identified by Spensley et al. (2009) had functional roles in the regulation of *LHY* expression had not previously been investigated. To discover whether they were sufficient to drive rhythmic expression, *luciferase* reporter constructs driven by multimers were designed and generated as described in Chapter 2. These multimers consisted of triple repeats of each motif isolated from its context in the *LHY* promoter, situated upstream of a minimal promoter providing a basal level of transcription. Multimers were also

generated to test the function of other isolated promoter motifs, including the G-box and 5A motifs. In addition, potential interactions between motifs were investigated by generating similar constructs using combinations of three different isolated motifs, such as a G-box flanked by two 5A motifs, outside of the context of the *LHY* promoter. Constructs to test the putative Elements 1-3 in combination with their neighbouring motifs were also generated.

However, when these constructs were transformed into plants, it became clear that the minimal *nos* promoter, required to drive a basal level of transcription of the constructs, was not functioning correctly. Due to time-constraints, these experiments therefore had to be put aside for future investigation. However, the binding of transcription factors to Elements 1, 2 and 3 was later examined in a Yeast One-Hybrid screen, as can be seen in Chapter 5.

3.3 - Discussion

3.3.1 - Summary of Conclusions

- The CT-rich region is essential for rhythmic expression of *LHY*.
- The CT-rich region is also required for activation of the promoter, but only when the distal promoter region is not present. Therefore the distal region of the promoter compensates for the effect of deleting the CT-rich region, suggesting that there are at least two nodes of activation on the promoter, but only one of them is required for expression.
- The presence of 5A motifs in the (-957/+1) *LHY* promoter is essential for its expression.

- The 5A motifs are required for transcriptional activation and act to enhance activation by the distal promoter region, either by assisting the binding of an activator or antagonising the action of a repressor in this region (Figure 3.5).

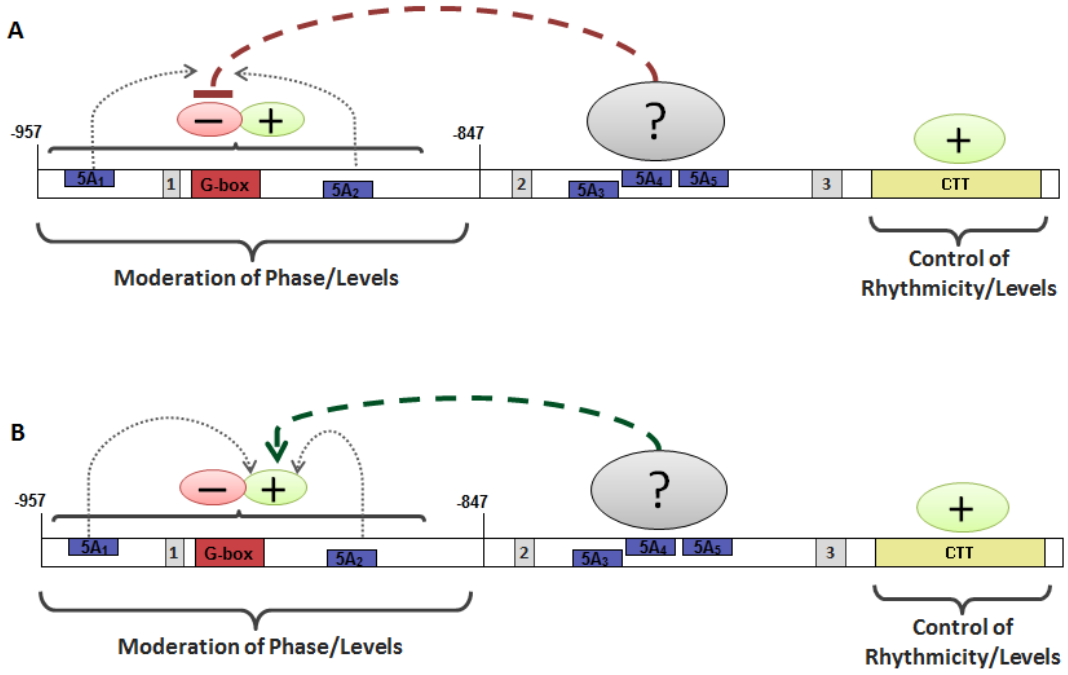


Figure 3.5: Specific promoter motifs mediate regulation of *LHY* transcription, with two possible mechanisms. The CT-rich region promotes activation of LHY:LUC reporter constructs, and is required for rhythmicity. Repressor(s) and activator(s) act on the distal promoter region to modulate expression. The proximal 5A motifs aid activation in the distal promoter region. The 5A motifs either A: antagonise the action of the repressor(s), or B: assist the action of the activator(s) in the distal promoter region.

3.3.2 - Discussion of Conclusions

We have proposed that there are at least two redundant regions of activation on the *LHY* promoter. This may indicate a possible route for multiple input signals to modify the expression of clock function, potentially allowing for greater flexibility in response to a changing environment. Multiple mechanisms of transcriptional

activation are likely to confer robustness to the clock, enabling rhythmic expression to persist under a wide range of conditions.

We can now assign a functional role to the CT-rich region of the *LHY* promoter: it comprises a binding site(s) for transcriptional activator(s), and is required for rhythmic expression of *LHY*. Not many promoter elements are known to be associated with circadian rhythmicity (Adams & Carre, 2011). However, it is not clear whether the CT-rich region mediates rhythmicity on its own, or whether it simply facilitates interaction between rhythmic transcription factors and the transcriptional machinery.

The CT-rich region resembles the $(GA)_n$ motif that is bound by *Arabidopsis* basic pentacysteine proteins (BPCs) (Meister et al., 2004). These proteins are thought to affect expression of genes in a wide range of processes, primarily those related to cell growth, development and patterning in multiple organs, including rosette leaves, hypocotyls, lateral roots and seeds (Monfared et al., 2011). The BPC proteins are also known to be capable of inducing conformational changes of promoters through cooperative binding, as shown by the action of BPC1 at the GA-rich sequence of the ovule developmental gene SEEDSTICK (STK) (de Folter et al., 2005). However, BPCs are not the only candidate transcription factors that could be targeting the CT-rich region of the *LHY* promoter. A similar repeated sequence was also identified in a later study as being over-represented in genes upregulated by induction of the circadian clock gene TIMING OF CAB1 (TOC1) (Gendron et al., 2012), suggesting that the CT-rich region may also be important for mediating regulation of *LHY* by

TOC1. Alternatively, this may suggest that TOC1 acts together with proteins binding GA motifs. These possibilities are investigated further in Chapter 4.

We were able to clarify the role of the 5A motifs as positively regulating *LHY* expression by interaction with (a) motif(s) in the distal promoter region. Within this distal region, there were two possible motifs that could have been targeted for this interaction: the G-box, and Element 1. The close proximity of these motifs allows for the possibility that both may be involved in this interaction. Further investigation, as described in Chapter 5, was required to determine if one or both of these could interact with the 5A motifs. A motif similar to the Evening Element (AAATATCT), termed EE-Like-expanded ((g/t/a)AA(g/t)ATC(g/t/c)) was subsequently identified by Huang et al. (2012) as being over-represented in promoters bound by TOC1. This motif is highly similar to a sequence overlapping Element 1 in reverse orientation in the *LHY* promoter (EE-Like-expanded: (T)AA(G)ATC; *LHY* promoter: CTACAAT). It is unclear whether such incidences of reverse orientated motifs within promoters are meaningful. However, the presence of this motif within the *LHY* promoter remains noteworthy, though further study is required to determine whether it has a role in the context of the *LHY* promoter.

The function of the putative regulatory motifs Elements 1-3 remains unknown, though the conservation across species of these sequences in the *LHY* promoter suggests a functional role. As described by Spensley et al. (2009) these elements do not contain any sequences known to regulate rhythmic expression, although they all bear some similarity to the CCAC sequence of the Morning Element (CCACAC) which is

associated with morning-expressed genes (Michael et al., 2008). Elements 1 and 3 contain the related SORLIP1 (GCCAC) sequence, which is known to be over-represented in the promoters of light-regulated genes, specifically phytochrome (PHYA) induced genes (Hudson & Quail, 2003), and Element 2 contains a GCCTC sequence highly similar to SORLIP1. The similarities to such light-regulated and morning-specific promoter elements suggests light-dependent regulatory roles for Elements 1-3, potentially providing additional sites of light signalling input to the *LHY* promoter beyond that of the G-box. However, the precise roles of these Elements in the regulation of *LHY* expression have yet to be determined. Although the multimers generated to address the question were unable to do so due to a problem with the minimal promoter sequence, these constructs are likely to be adapted for future experiments to investigate the regulatory roles of Elements 1-3.

CHAPTER 4

TOC1 Regulates *LHY* Expression through Interactions with Specific Promoter Motifs

4.1 - Introduction

As described in Chapter 3, the *LHY* promoter was known to comprise multiple evolutionarily conserved sequence motifs, including the G-box, 5A motifs and the CT-rich region. The G-box had been proposed by Spensley et al. (2009) to moderate the phase and level of *LHY* expression through both transcriptional activation and repression. We showed in Chapter 3 that both the CT-rich region and 5A motifs mediate activation, and that the CT-rich region is essential for rhythmicity. We also found that the CT-rich region and distal promoter region act redundantly to activate expression, and proposed that the proximal 5A motifs promote expression of *LHY* through the prevention of repression by the distal promoter region, possibly through the G-box (Figure 3.5).

Although it was not known which transcription factors might be targeting these motifs to regulate *LHY* expression, the other circadian clock proteins were obvious candidates. As described in the Introduction (Figure 1.1), a number of different transcription factors have been identified as part of the multiple feedback loops regulating *LHY*. These include APRR9, APRR7, APPR5, APRR3 and APRR1 (TOC1), all negative regulators of *LHY* that are expressed sequentially at intervals of approximately 2 hours (Matsushika et al., 2000, Gendron et al., 2012, Huang et al., 2012). PRRs 9, 7 and 5 are known to bind the *LHY* promoter around the G-box (Nakamichi et al., 2010). At the start of this project, TOC1 was not known to bind

DNA, although it was recently shown to be capable of directly binding DNA *in vitro* and to associate with the distal region of the *LHY* promoter *in planta* (Gendron et al., 2012).

It should also be noted that despite TOC1's role as a transcriptional repressor of *LHY*, the complexity of interactions between negative feedback loops in the clock results in the net genetic effect of TOC1 being to reactivate expression of *LHY* at dawn (Alabadi et al., 2001, Más et al., 2003a, Huang et al., 2012). TOC1 is therefore both directly targeting the *LHY* promoter to repress transcription, and indirectly causing activation at the *LHY* promoter through its effects on other transcriptional regulators.

Largely due to this complexity of regulation, the precise mechanism by which TOC1 regulates *LHY* expression is unknown; at the start of this project it was not even known to bind DNA directly. We therefore wanted to identify *LHY* promoter motifs mediating the effects of TOC1, and hence begin to unravel the mechanism by which TOC1 regulates the expression of *LHY*.

Our experimental aims were therefore as follows:

- Investigate whether the effects of TOC1 are mediated by the proximal and/or distal *LHY* promoter regions (Section 4.2.1)
- Which promoter motifs mediate the effects of TOC1? (Section 4.2.2)
- Identify which promoter motifs are bound by TOC1 (Section 4.2.3)
- Identify regulatory effects of TOC1 (Section 4.2.4)

4.2 - Results

4.2.1 - The Proximal Region of the *LHY* Promoter Mediates Activating Effects of TOC1

In order to investigate which promoter regions were targeted by TOC1, full-length (-957/+1) and truncated (-847/+1) *LHY* promoter constructs were transformed into TOC1 RNAi plants (Más et al., 2003a) which had had their *cab:luc* reporter constructs crossed out, and TOC1 overexpressor (TOC1ox WS) plant lines (from Dr László Kozma-Bognár). The TOC1ox plants displayed a long-period clock phenotype (Figure 4.2), consistent with published data in other plant lines with increased TOC1 expression (Más et al., 2003a). Expression of LHY:LUC reporter constructs was analysed as described in Section 3.2, with 7-day old plants entrained to 12L:12D white light conditions, and imaged under red-light for 2 days in 12L:12D, followed by 5 days in constant light.

Expression in constant light of both full-length and truncated LHY:LUC constructs was significantly reduced in TOC1 RNAi plants (Figure 4.1A). In contrast, there was significant increase in mean luminescence for both constructs in TOC1ox plants (Figure 4.1B). This suggested that TOC1 promotes expression of both the full-length and truncated LHY:LUC constructs. Therefore, we propose that the proximal (-847/+1) region of the *LHY* promoter mediates transcriptional activation downstream of TOC1.

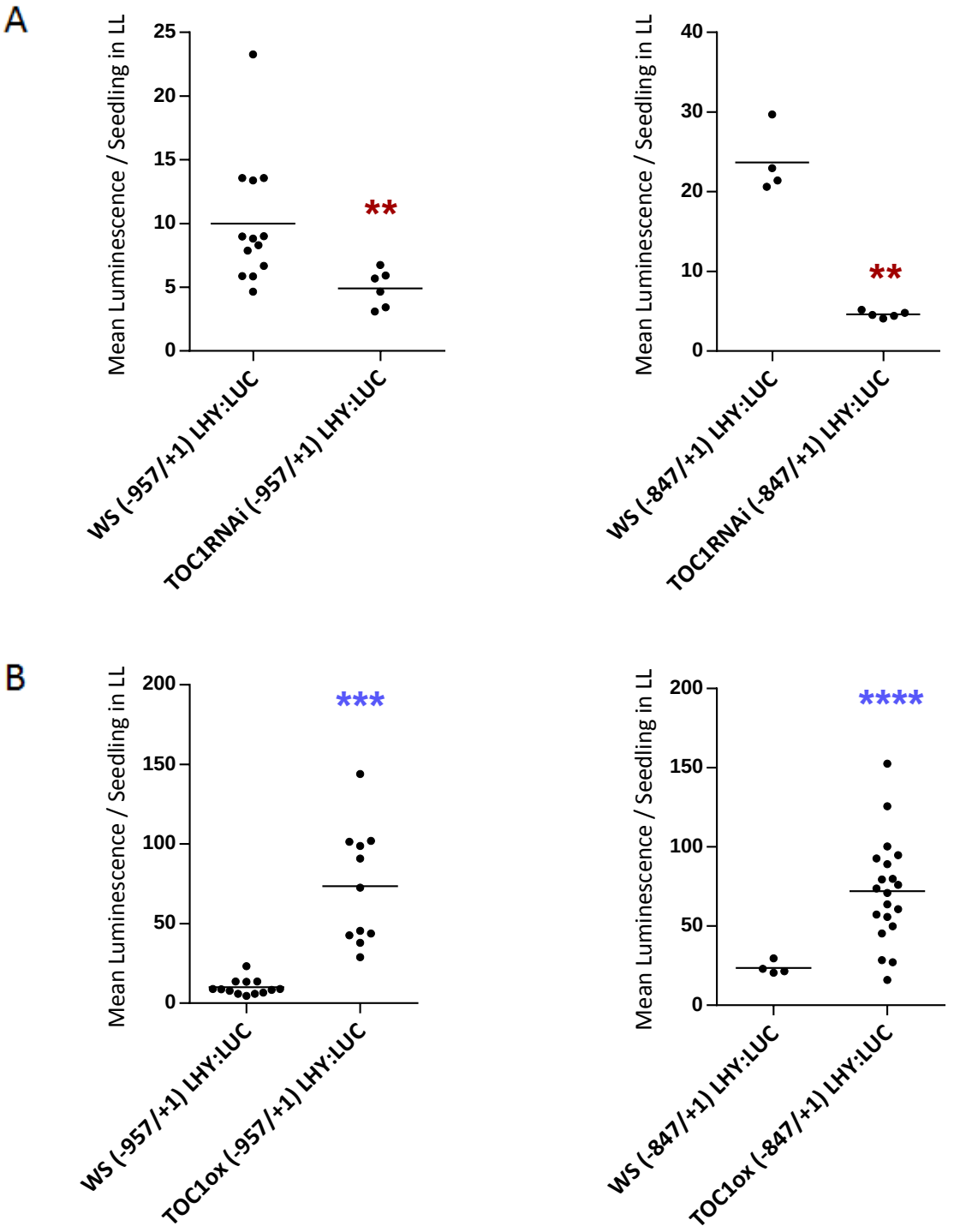


Figure 4.1: Effects of inhibition and overexpression of TOC1 on expression levels of full-length and truncated LHY:LUC reporter constructs. A: (-957/+1) LHY:LUC in wild-type (WS), TOC1 RNAi and TOC1ox plants. B: (-847/+1) LHY:LUC in wild-type (WS), TOC1 RNAi and TOC1ox plants. Expression levels per seedling for each transgenic line were averaged over a period of 72 hours in constant red light (after 7 days entrainment to 12L:12D). Each data point represents the mean expression level from one transgenic line. Mean levels for each reporter construct are shown by horizontal lines. p-values from T-tests are shown (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; NS = not significant). Red stars indicate comparisons between TOC1 RNAi plants and the WS control, and blue stars comparison between TOC1ox plants and the WS control.

Although expression was significantly reduced for the full-length LHY:LUC construct in TOC1 RNAi plants, rhythmic expression was visible under diurnal conditions (Figure 4.2A,C). These diurnal rhythms were of reduced amplitude but unaltered phase of expression (Figure 4.2C). In TOC1ox plants, full-length and truncated constructs displayed increased amplitude and unaltered phase of expression under light/dark conditions (Figure 4.2B,E). These results indicate that TOC1 mediates diurnal as well as circadian activation, but suggest that TOC1 is not involved in regulating the timing of *LHY* re-activation at dawn. This acute activation at dawn was not observed with the truncated LHY:LUC construct in TOC1 RNAi plants (Figure 4.2D), suggesting that the distal promoter region mediates this TOC1-independent response to dawn.

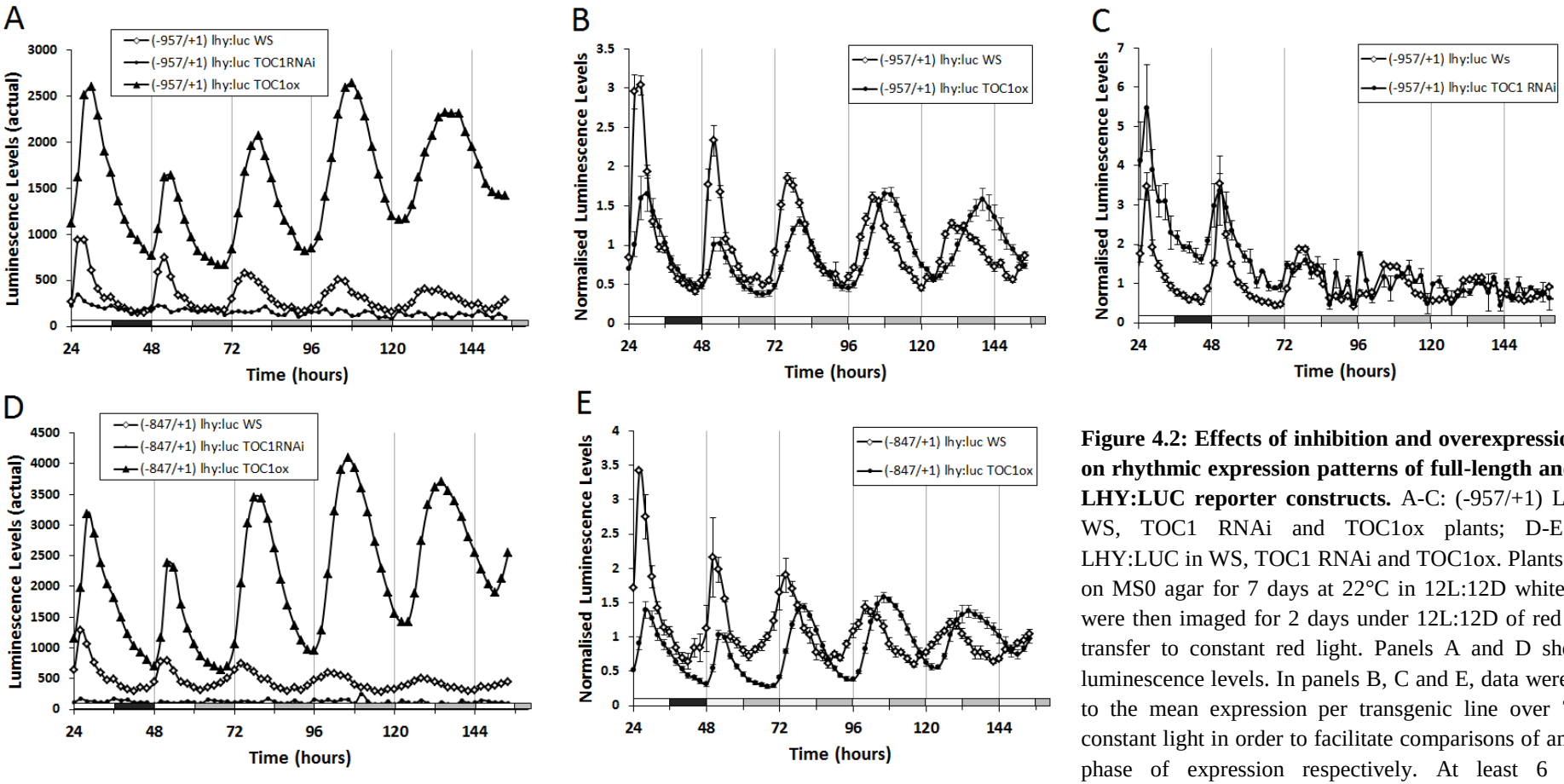


Figure 4.2: Effects of inhibition and overexpression of TOC1 on rhythmic expression patterns of full-length and truncated LHY:LUC reporter constructs. A-C: (-957/+1) LHY:LUC in WS, TOC1 RNAi and TOC1ox plants; D-E: (-847/+1) LHY:LUC in WS, TOC1 RNAi and TOC1ox. Plants were grown on MS0 agar for 7 days at 22°C in 12L:12D white light. They were then imaged for 2 days under 12L:12D of red light before transfer to constant red light. Panels A and D show absolute luminescence levels. In panels B, C and E, data were normalised to the mean expression per transgenic line over 72 hours of constant light in order to facilitate comparisons of amplitude and phase of expression respectively. At least 6 independent transgenic lines were analysed per construct. The data shown are averages from these multiple lines. Error bars indicate the Standard Error of the Mean (SEM). The expression pattern of (-847/+1) LHY:LUC is not shown for TOC1 RNAi plants, as its expression did not rise above background levels.

4.2.2 - The CT-rich Region May Facilitate Recruitment of TOC1 to the Proximal Promoter

The CT-rich region had been identified in Section 3.2.2 as essential for expression of the proximal promoter region. As shown in Figure 4.3A, inhibition of *TOC1* expression also effectively abolished transcription of the proximal promoter region, an effect similar to that of deleting the CT-rich region. This suggested that TOC1 might act either directly through the CT-rich region or in concert with transcription factors binding at the CT-rich region.

Interestingly, TOC1 overexpression restored both expression and rhythmicity to the -847 CTRm construct (Figure 4.3A,B). Therefore, we suggest that TOC1 does not directly target the CT-rich region but that this region acts to facilitate recruitment of TOC1 to the proximal promoter, and that its absence can be compensated for by other transcription factors when TOC1 is overexpressed.

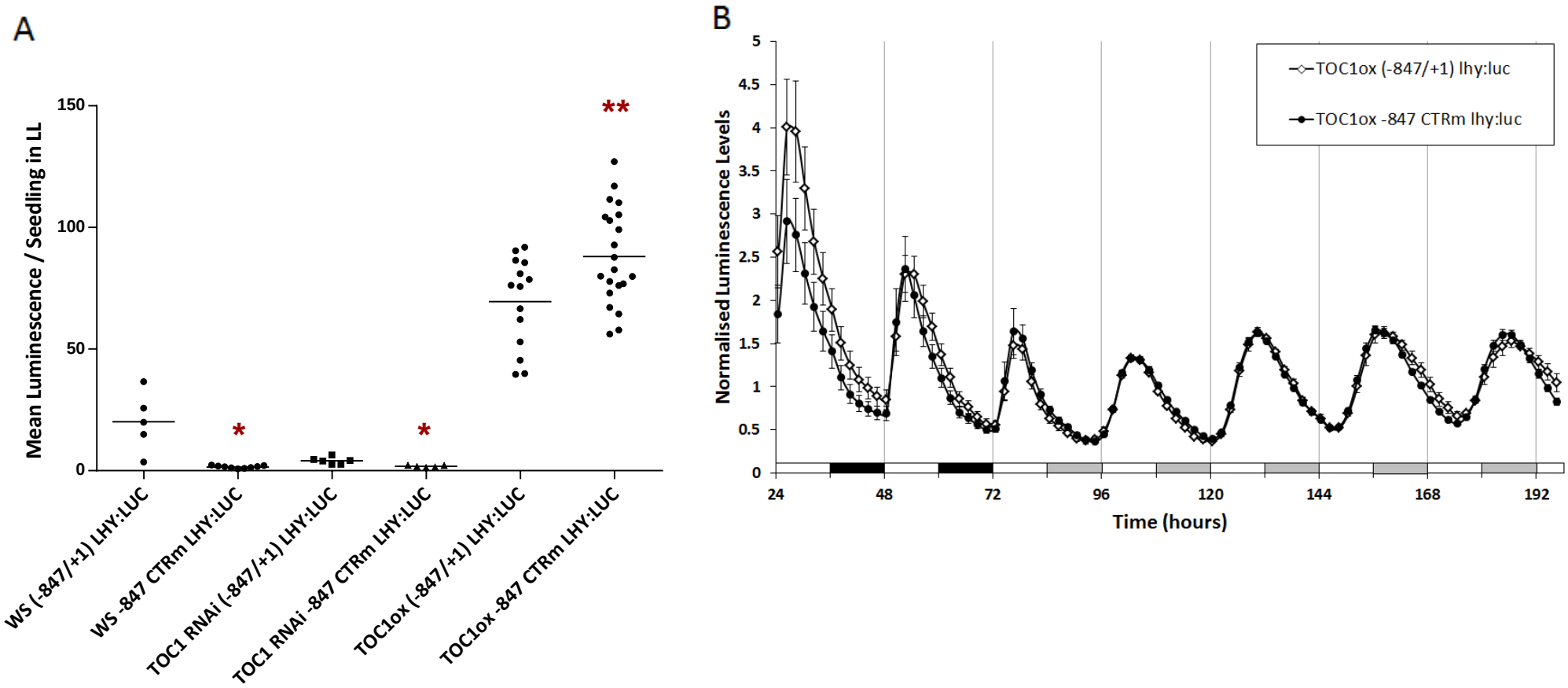


Figure 4.3: TOC1 overexpression can compensate for the effect of deleting the CT-rich region (CTR). A: Expression levels of (-847/+1) LHY:LUC and -847 CTRm LHY:LUC reporter constructs, averaged over a period of 72 hours in constant red light. Red stars indicate p-values from T-tests comparing -847 CTRm LHY:LUC and (-847/+1) LHY:LUC in either WS, TOC1 RNAi or TOC1ox plants (*p<0.05; **p<0.01; NS = not significant). B: Temporal expression patterns of the (-847/+1) LHY:LUC and -847 CTRm LHY:LUC reporter constructs in TOC1ox. Plants were grown on MS0 agar for 7 days at 22°C in 12L:12D white light, then imaged for 3 days in 12L:12D red light followed by 5 days constant red light. Data for each construct was normalised to mean expression level in constant light and averaged across transgenic lines. Error bars indicate standard errors.

4.2.3 - TOC1 Binds the *LHY* Promoter Around the G-box/Element 1

To establish where TOC1 binds on the *LHY* promoter, ChIP was performed on samples of the TOC1 Minigene plant line, a *toc1* mutant line containing YFP-tagged TOC1 under the control of a *TOC1* promoter (Más et al., 2003a). Q-PCR analysis was then performed using primer sets distributed along the *LHY* promoter from upstream of the G-box to the start of the CT-rich region (Methods, Table 2.1). The CT-rich region itself was not tested due to the difficulty of designing primers to such a large repetitive sequence. However, the most downstream primer pair tested (primer set 4) maps to immediately upstream (within 2 basepairs) of the start of the CT-rich region, and therefore any binding by TOC1 in the CT-rich region should be detectable as an increase in recovery of Input DNA corresponding to this adjacent promoter region.

As can be seen in Figure 4.4, TOC1 did not appear to be highly associated with the promoter region adjacent to the CT-rich region (primer set 4). This region displayed the lowest percentage recovery of all the promoter regions tested, suggesting that TOC1 does not bind the CT-rich region. In contrast, TOC1 was found to associate strongly with the *LHY* promoter region containing the G-box and novel Element 1 (primer set 2), with a two-fold higher percentage recovery of Input than in adjacent regions (primer sets 1 and 3). TOC1 therefore appears to be binding close to the G-box in the distal region of the *LHY* promoter.

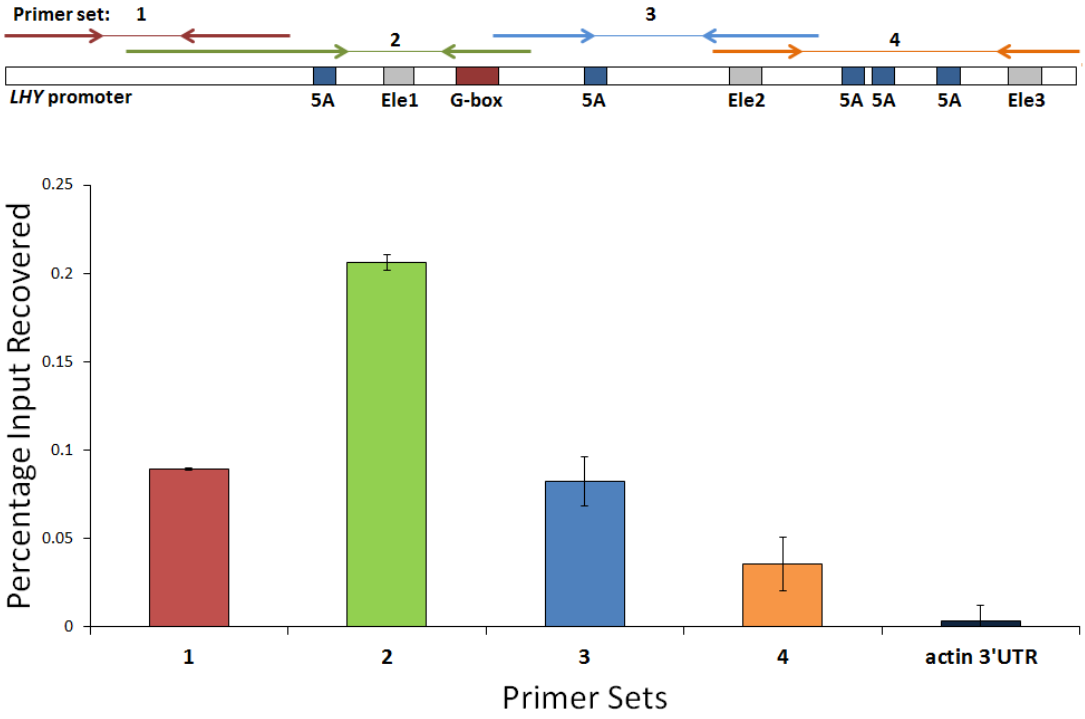


Figure 4.4: TOC1 associates with the *LHY* promoter around the G-box. In order to assay binding of TOC1 at different locations along the *LHY* promoter, TMG plants (*pTOC1::TOC1:YFP* in *toc1-2* background) were grown under 12L:12D conditions at 22°C for 14 days then transferred to constant light. Tissue was harvested at ZT16 in the first LL cycle. Chromatin Immunoprecipitation (ChIP) experiments were carried out using an antibody to YFP. Enrichment for *LHY* promoter sequences was tested by Q-PCR. Amplification of actin 3'UTR sequences was used as a negative control for ChIP enrichment (black). Primers were designed to amplify overlapping adjacent regions in the -957/-754 region of the *LHY* promoter, as shown by the diagram above the graph. Enrichment of sequences was calculated as the percentage of Input recovered. Data shown are means of technical replicates. Error bars indicate standard deviations.

4.2.4 - The G-box Mediates Regulation by TOC1 on *LHY* Expression

Luciferase assays of full-length *LHY* promoter constructs with altered G-box flanking nucleotides (-957 Wk Cl II LHY:LUC) were performed in wild-type (WS) and TOC1 RNAi plants to investigate the potential regulatory role of TOC1 at the G-box.

Spensley et al. (2009) found that the (-957/+1) LHY:LUC construct containing this G-box mutation drove luminescence rhythms with a reduced amplitude of expression. In addition, the transcriptional peak was broader in constant light than in wild-type plants entrained to short days of 8L:16D. Similar results were obtained under 12L:12D cycles (Figure 4.5A).

In TOC1 RNAi plants entrained to 12L:12D then transferred to constant light, this effect of the G-box mutation was abolished and there was no significant difference in the transcriptional peak or amplitude of expression between (-957/+1) LHY:LUC and -957 Wk Cl II LHY:LUC constructs (Figure 4.5B). This result was consistent with regulation of rhythmic expression by the G-box being dependent on the presence of TOC1 in the system.

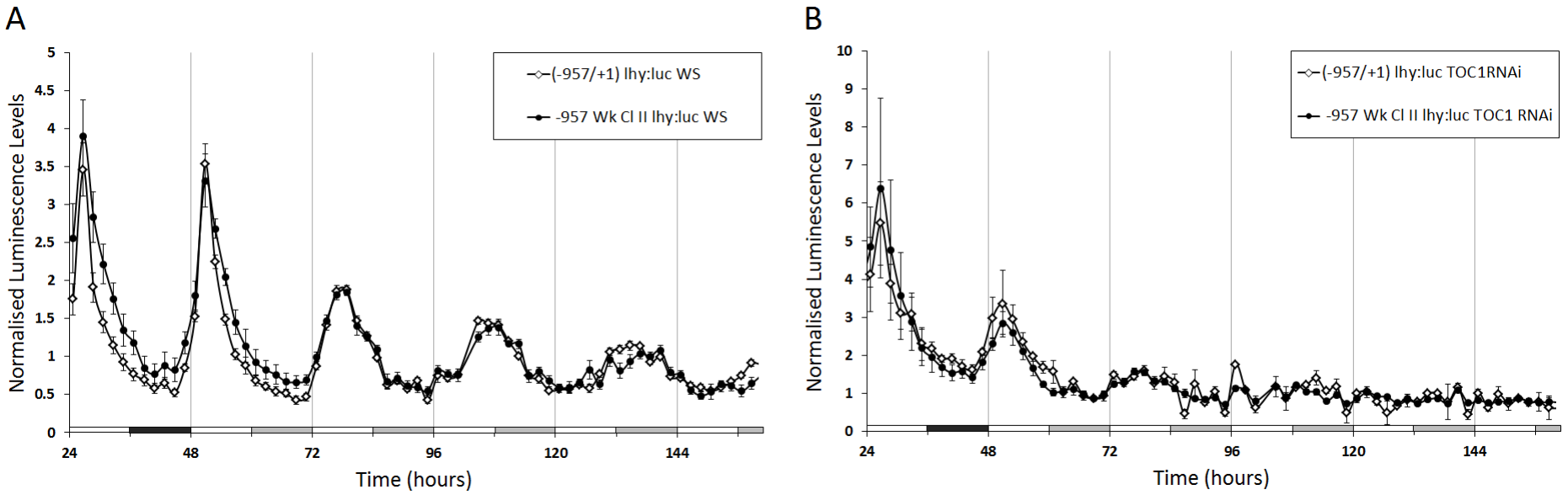


Figure 4.5: Mutation of the G-box motif does not alter the effect of TOC1 on rhythmic patterns of *LHY* transcription. Plants were grown on MS0 agar for 7 days at 22°C in 12L:12D white light. They were imaged for 2 days in 12L:12D red light then for 5 days of constant red light. Temporal patterns of luminescence were normalised to the mean expression level in constant light then averaged across independent transgenic lines.

In TOC1 RNAi plants, the mean expression level of the (-957/+1) LHY:LUC construct decreases significantly, indicating that TOC1 activates expression of this promoter construct. In wild-type plants, mutation of the G-box resulted in minor (Spensley et al., 2009) or insignificant (Figure 4.6) effects on (-957/+1) LHY:LUC expression levels. However, expression increases in the TOC1 RNAi background when the G-box is mutated (Figure 4.6). This suggests that the G-box mutation disrupts binding of a transcriptional repressor, and that these effects are only revealed when TOC1 activity is reduced by RNAi.

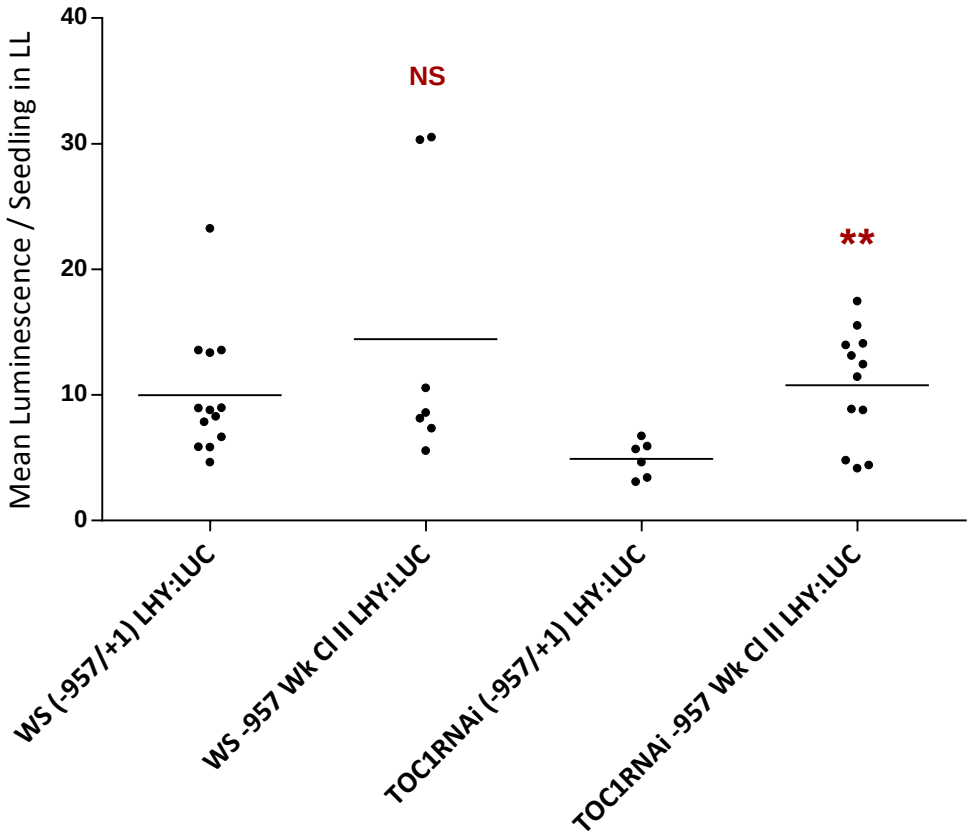


Figure 4.6: Mutation of the G-box increases expression of the LHY:LUC reporter gene in TOC1 RNAi plants. Plants were grown on MS0 agar for 7 days at 22°C in 12L:12D white light. They were imaged for 1 day in 12L:12D red light then for 5 days of constant red light. Expression levels were averaged over a period of 72 hours of constant red light, and averaged over multiple transgenic lines. Red stars indicate p-values from T-tests comparing (-957/+1) LHY:LUC and -957 Wk Cl II LHY:LUC in either WS or TOC1 RNAi plants (**p<0.01; NS = not significant).

We therefore propose that TOC1 is indirectly activating expression at the G-box by antagonising the action of another transcriptional repressor. This might occur through direct action, by antagonising the binding of this repressor to the G-box. Or it could be that TOC1 represses the expression of this other transcriptional repressor of *LHY*.

The former hypothesis appears unlikely; since TOC1 is now known to directly repress *LHY* expression, it would seem counterproductive for TOC1 to simultaneously antagonise the binding of another repressor. However, the hypothesis that TOC1 might repress the expression of another repressor allows for a time delay between the action of TOC1 and the unknown repressor. In this scenario, TOC1 would be simultaneously repressing the expression of both *LHY* and the other repressor. Therefore, when TOC1 protein levels have reduced sufficiently to allow expression of both, there would be a period of time during which *LHY* could be expressed before the repressor protein could accumulate enough to inhibit *LHY* expression.

4.3 - Discussion

4.3.1 - Summary of Conclusions

TOC1 is required for wild-type expression of both full-length and truncated constructs, and appears essential for activation of transcription by the proximal promoter region. However, TOC1 RNAi does not abolish diurnal expression from the full-length construct, suggesting TOC1-independent light-induced activation of transcription by the distal region of the promoter. As the effect of deleting the CT-

rich region on truncated LHY:LUC expression bears a striking similarity to the lack of detectable expression seen in TOC1 RNAi plants, we suggest that TOC1 interacts with transcription factors within the CT-rich region to mediate activation of transcription. TOC1 was found to bind around the G-box and novel Element 1, suggesting that it may be directly regulating *LHY* expression through either of these motifs. TOC1 was also found to be indirectly activating *LHY* expression through the G-box, potentially by inhibition of another G-box-binding transcriptional repressor.

4.3.2 - TOC1 Both Indirectly Activates and Directly Represses *LHY* Expression at the G-box

As described in the Introduction, at the start of this project TOC1 was not known to contain any DNA-binding domains, and the mechanism by which it regulated *LHY* expression was unknown. Gendron et al. (2012) subsequently published ChIP data also showing that TOC1 binds a region of the *LHY* promoter corresponding to the distal promoter region. However, they did not test other areas of this functional promoter so could not rule out binding elsewhere. In addition, Huang et al. (2012) showed TOC1 binding at the *LHY* promoter through ChIP-seq, upstream of the transcriptional start site (which is flanked by the CT-rich region).

Through anti-GFP ChIP experiments on a TMG line containing YFP-tagged TOC1, we found that TOC1's binding profile along the *LHY* promoter focussed specifically on a small region within the distal *LHY* promoter containing a 5A motif, the G-box and the putative regulatory motif Element 1. Of these, only the G-box and Element 1 motifs are not found elsewhere on the promoter and are thus potential binding targets for TOC1. Since Element 1 and the G-box are situated just 6 basepairs apart on the

LHY promoter, it is plausible that TOC1 may be interacting with both. Element 1 does not correspond to the current predicted binding motifs of TOC1. Therefore, further investigation is required to determine whether Element 1 may be a binding target of TOC1.

The G-box is known to be preferentially enriched in the promoters of genes upregulated by TOC1 (Gendron et al., 2012), so is a likely target in this promoter. However, it was also shown by Huang et al. (2012) and Gendron et al. (2012) that the role of TOC1 on *LHY* expression is repressive. The apparent activation by TOC1 seen here in TOC1ox and TOC1 RNAi plants is in line with previous publications (Más et al., 2003a) and is an indirect effect of TOC1's action on other regulators of *LHY*. TOC1 is therefore likely to be targeting the G-box for direct repression of *LHY*.

We also proposed a secondary role for the G-box in mediating indirect activation by TOC1, such that the G-box is targeted by a transcriptional repressor whose action is antagonised by TOC1. As described in Section 4.2.4, we suggested that while TOC1 is regulating *LHY* expression through direct binding around the G-box, it could also be repressing the expression of another transcription factor. This would allow for a delay between the relief of repression of *LHY* by TOC1, and the sufficient build-up of the other transcription factor to levels where it can once more repress *LHY* expression after dawn. A likely candidate for this TOC1-regulated repressor is PRR9, which is known to bind and repress the *LHY* promoter around the G-box after dawn (Nakamichi et al., 2010) and whose expression is negatively regulated by TOC1 (Makino et al., 2002).

4.3.3 - Regulation by TOC1 in the Proximal Region of the *LHY* Promoter

Since the CT-rich region and TOC1 are both required for expression of the proximal promoter, we suggested that the CT-rich region may facilitate recruitment of TOC1 to the promoter. The CT-rich region is unlikely to mediate direct binding of TOC1 since ChIP experiments did not show TOC1 binding to a region immediately upstream of the CT-rich region. This correlates with ChIP-seq data published by Huang et al. (2012) showing TOC1 binding the *LHY* promoter upstream of the transcriptional start site, and therefore upstream of the CT-rich region.

Induction of TOC1 expression was shown to result both in activation and repression of gene expression (Gendron et al., 2012). The G-box and a GA-repeat motif, highly similar to the CT-rich region, were enriched in the promoters of genes that were upregulated by TOC1, whereas the TBS motif was enriched in genes that were downregulated. In accordance with this, we showed that the CT-rich region mediates transcriptional activation of *LHY*. However, this does not necessarily suggest a directly activating role for TOC1, since truncating the *LHY* promoter to remove the distal promoter region may have removed binding sites for TOC1 cofactors that would normally mediate transcriptional repression.

Alternatively, TOC1-dependent activation of the proximal promoter region could occur indirectly, through transcriptional regulation of activators or repressors that target this region. Similarly, any recruitment of TOC1 or TOC1-regulated transcription factors to this promoter region could displace other transcriptional regulators at the CT-rich region. As discussed in Section 3.3, the CT-rich region is predicted to be specifically targeted by the Basic Pentacysteine (BPC) family of

transcription factors. The BPCs have been shown to bind *in vitro* to GA repeat sequences similar to the CT-rich region sequence, and to have complex, overlapping and antagonistic functions in the regulation of multiple developmental processes in plants, including cell growth, ovule development and leaf morphology (Monfared et al., 2011).

CHAPTER 5

Transcription Factors Binding the *LHY* Promoter

5.1 - Introduction

Chapters 3 and 4 assigned regulatory roles to several of the evolutionarily conserved sequence motifs within the *LHY* promoter, and identified the G-box as mediating both direct and indirect regulatory effects of TOC1. The CT-rich region is also likely to be involved in mediating regulation by TOC1. However, the role of the novel Elements 1, 2 and 3 in the regulation of *LHY* expression remained unknown.

At the start of this project, little was known about the direct regulation of *LHY* expression by transcription factors. Although known motifs such as the G-box allowed identification of families of transcription factors likely to bind the *LHY* promoter, the binding partners of novel sequence motifs had yet to be identified. As described in Chapter 1, a number of circadian clock proteins were thought to regulate the expression of *LHY*, including CCA1, LUX, TOC1, PRR9, PRR7, PRR5 and LHY itself (Wang & Tobin, 1998, Hazen et al., 2005, Makino et al., 2002, Nakamichi et al., 2005). CCA1 was known to interact *in vitro* with a AAA($^{A}/_{C}$)AATCTA sequence (Wang et al., 1997) which is present in the *LHY* promoter downstream of the conserved region (626 basepairs upstream of the translational start site). However, whether CCA1 could interact with this sequence in the context of the *LHY* promoter remained unknown. Similarly, direct binding of other clock proteins to the *LHY* promoter had yet to be established. It was subsequently discovered that PRR9, PRR7 and PRR5 bind the *LHY* promoter in the region of the G-box *in vivo* (Nakamichi et al., 2010).

Non-circadian transcription factors were also known to regulate *LHY* expression. The floral repressor FLC was known to directly bind to the *LHY* promoter *in planta* (Spensley et al., 2009). This binding of FLC was proposed to occur through the 5A motifs of the *LHY* promoter, since they resembled the known FLC binding site on the *SOC1* promoter. Light-dependent regulation of *LHY* expression was thought to occur through the phytochrome-interacting transcription factor PIF3, which can bind *in vitro* to a G-box-containing fragment of the *LHY* promoter both alone and in complex with PhyB (Ni et al., 1998, Martínez-García et al., 2000). However, this interaction had not been shown to occur *in vivo* or in the wider context of the *LHY* promoter.

Therefore, although transcription factors had been identified as candidates for the regulation of *LHY* expression, very few had been confirmed as binding the *LHY* promoter. We therefore wanted to identify transcription factors capable of binding the promoter and to map their binding to specific regulatory elements within the promoter.

Aims

- Identify transcription factors capable of binding the *LHY* promoter (Section 5.2.1)
- Determine which promoter motifs these transcription factors bind (Section 5.2.2)

5.2 - Results

5.2.1 - Identifying Transcription Factors Binding the *LHY* Promoter

5.2.1.1 - Yeast One-Hybrid

Transcription factors able to bind the *LHY* promoter were identified with a Yeast One-Hybrid screen of a collection of 1181 *Arabidopsis* transcription factors (constructed from the REGIA clone library (Paz-Ares, 2002) by the PRESTA group, Warwick HRI). A Yeast One-Hybrid screen consists of analysing the growth under selective conditions of multiple independent strains of *S. cerevisiae* each containing two plasmids. The first plasmid encodes a variable 'prey' transcription factor fused to a yeast activation domain (AD). The second plasmid contains a 'bait' DNA sequence upstream of a yeast promoter and a reporter gene. When the prey protein binds to the bait sequence, it brings the AD into proximity with the yeast promoter, causing activation of the reporter gene (Figure 5.1A). Typically, expression of the reporter gene enables survival of the yeast under specific conditions. In our screen, yeast strains unable to synthesise the essential amino acid Histidine were grown on selective media lacking it. Only yeast containing a successful interaction between prey and bait sequence, and hence expressing the Histidine reporter gene, could survive under these conditions.

pDEST22 (*trp* selection) vectors encoding individual transcription factors tagged with a yeast activation domain were transformed into AH109 haploid mating yeast in groups of 12. Two arrangements of the transcription factor library were tested simultaneously, each pooled into groups containing different combinations of transcription factors. Bait sequence was constructed by amplifying a fragment of the *LHY* promoter from 957 to 747 basepairs upstream of the translational start site. This

bait sequence did not include the CT-rich region due to size constraints but included all other known evolutionarily conserved promoter motifs described in Chapter 3. The promoter fragment was cloned into a pHIS3LEU2 vector (*leu* selection for presence of plasmid) and transformed into a haploid Y187 yeast strain. This haploid pLHY Y187 strain was independently mated with each pool of haploid TF+AD AH109 strains (Figure 5.1B).

Growth of the subsequent diploid yeast on SD-leu-trp-his (SD-LTH) media should only occur if at least one of the 12 AD-tagged prey transcription factors in that pool were able to bind to the bait *LHY* promoter fragment. However, the HIS3 reporter gene can still be expressed at a low level in the absence of a prey-bait interaction. This auto-activation can occur through binding of native yeast proteins to the bait sequence and can mask prey-bait interactions. The addition of 3-Amino-1,2,4-triazole (3-AT), a competitive inhibitor of Histidine, to growth media removes the effects of this auto-activation. Concentrations of 3AT were determined for each bait construct. Since a higher prey-bait binding affinity will result in increased expression of Histidine, the screen was performed at a range of inhibitor concentrations (0mM, 50mM and 100mM 3AT) to assess the strength of prey-bait binding.

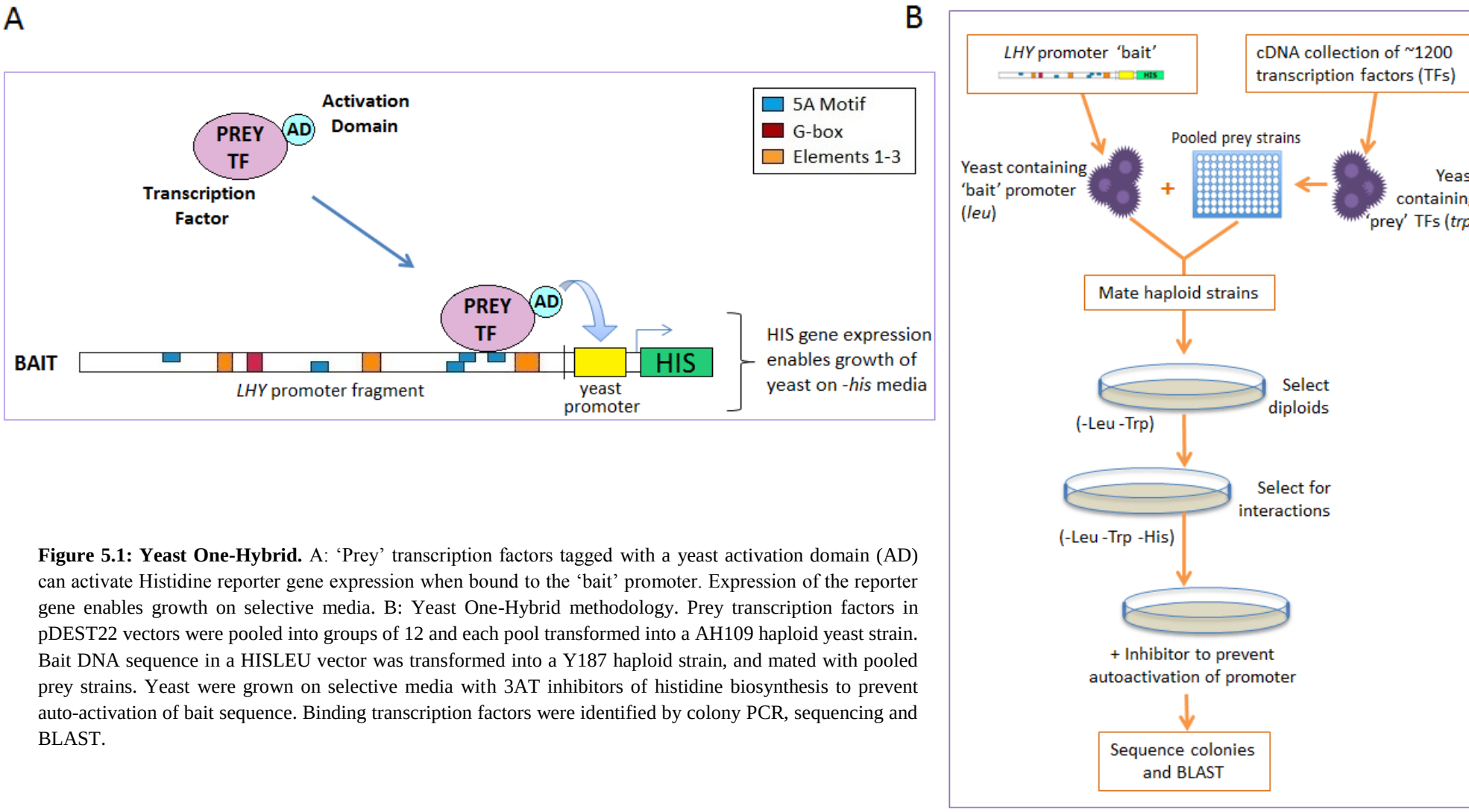


Figure 5.1: Yeast One-Hybrid. A: 'Prey' transcription factors tagged with a yeast activation domain (AD) can activate Histidine reporter gene expression when bound to the 'bait' promoter. Expression of the reporter gene enables growth on selective media. B: Yeast One-Hybrid methodology. Prey transcription factors in pDEST22 vectors were pooled into groups of 12 and each pool transformed into a AH109 haploid yeast strain. Bait DNA sequence in a HISLEU vector was transformed into a Y187 haploid strain, and mated with pooled prey strains. Yeast were grown on selective media with 3AT inhibitors of histidine biosynthesis to prevent auto-activation of bait sequence. Binding transcription factors were identified by colony PCR, sequencing and BLAST.

Five yeast colonies were sequenced from every pool of mated transcription factors that was able to survive on SD-LTH + 50mM or 100mM 3AT media. Transcription factors present in the yeast were identified using NCBI BLAST (Basic Local Alignment Search Tool) (Johnson et al., 2008). Each yeast colony should contain a single prey vector from the pool of 12 possible transcription factors. Therefore, transcription factors capable of binding the *LHY* promoter were identified as those present in the surviving yeast in either of the parallel screens.

Fifteen transcription factors were identified as binding to the -957/-754 *LHY* promoter fragment from the Yeast One-Hybrid screen, twelve from one library and three from both.

5.2.1.2 - Yeast One-Hybrid Results

Many of the transcription factors identified as binding the *LHY* promoter from Yeast One-Hybrid were re-tested individually in yeast to confirm their interaction. The results of these tests and of binding controls in all subsequent yeast assays involving these transcription factors are shown in Table 5.1.

Initial tests for confirmation of binding were carried out on the strongest binders, i.e. those transcription factors found to bind the promoter at the highest level of inhibitors (100mM 3AT). These were: PIF7 (gene model 2), NAM, ABF4, EEL, DPBF2, IAA2 and GBF1. Because of the number of ABA-related proteins strongly binding the promoter, ABF3 was also included in these tests. For transcription factors detected only on media containing lower levels of inhibitors, binding was individually tested in later assays (see Section 5.2.2).

Table 5.1: Yeast One-Hybrid results and confirmation of interactions. A: Binding of transcription factors (TFs) to the *LHY* promoter as identified from the initial Y1H screen, with the maximum 3AT inhibitor level of this binding. Growth at 100mM gives greater confidence that interaction is not a false positive. Library 1 and 2 represent biological replicates for the Y1H screen. PIF7 (5'tr) is a splice variant of PIF7 containing a 5'truncation. PIF7 (full) is re-amplified full-length PIF7 corresponding to gene model 1. B: Individual mating (of haploid yeast) and sequential transformation (of diploid yeast) tests were subsequently performed for the strongest binders to assess suitability for future assays. C: Biological replicates for these interactions are shown as the results of binding controls from future assays (Section 5.2.2 and Chapter 6).

		A: Y1H (pooled libraries) Chapter 5			**B: Individual TF Testing**		**C: Future Assays**	
							Section 5.2.2	Chapter 6
	TF Name	**Library 1**	**Library 2**	**Max. Inhibitors**	**mating**	**seq. transf.**	**mating**	**seq. transf.**
AT5G61270	**PIF7**	yes (5'tr)	yes (5'tr)	100mM	yes (5'tr)	yes (5'tr)	no (full)	no (full)
AT3G12910	**NAM**	yes	yes	100mM	yes	yes	yes	yes
AT3G19290	**ABF4**	yes	yes	100mM	no	weak	no	-
AT2G41070	**EEL**	no	yes	100mM	no	weak	no	-
AT3G44460	**DPBF2**	no	yes	100mM	no	weak	no	-
AT3G23030	**IAA2**	no	yes	100mM	no	yes	no	-
AT4G36730	**GBF1**	no	yes	100mM	yes	yes	yes	-
AT4G34000	**ABF3**	yes	no	50mM	yes	yes	yes	yes
AT4G37750	**ANT**	yes	no	50mM	-	-	yes	-
AT2G21230	**AtbZIP30**	yes	no	50mM	-	-	yes	-
AT1G06850	**AtbZIP52**	yes	no	50mM	-	-	-	-
AT4G38900	**AtbZIP29**	no	yes	50mM	-	-	yes	-
AT4G18110	**AT4G18110**	no	yes	50mM	-	-	no	-
AT3G61180	**AT3G61180**	no	yes	50mM	-	-	no	-
AT4G09180	**FBH2**	no	yes	50mM	-	-	no	-

Binding was tested in two ways: (1) using individual mating assays with two haploid strains, one containing the bait promoter and the other containing a single transcription factor, and (2) using Li-Ac transformation-based assays, where promoter and transcription factor were sequentially transformed into a single, diploid, yeast strain. These techniques can give different levels of sensitivity to an assay and hence produce different results; the reasons behind this are not fully understood.

The majority of transcription factors identified through Yeast One-Hybrid were found to bind again in at least one of the independent replicates. The exceptions were FBH2 and the two zinc finger proteins (AT4G18110 and AT3G61180). In addition, AtbZIP52 was not re-tested due to difficulties in culturing yeast from the glycerol stock. All of these were initially identified as weakly binding the promoter, as they were only seen at the lower (50mM) concentration of inhibitors. Weak binding could explain why they were not found to bind in later assays, although it is still possible that they could have been false positives in the initial screen. Further testing is therefore required to confirm these interactions.

Predominant amongst the identified proteins were those known to be involved in flowering, light signalling and hormone and stress responses. However, not all could be related to a known function; of some, in particular the zinc finger proteins (AT3G61180, AT4G18110) and the NAM-family protein (AT3G12910), little was known beyond basic structural predictions. Therefore, to gain information on the potential functions of these transcription factors, publicly available protein

interaction and gene expression data was examined and compiled into Tables 5.2 to 5.4.

Functional roles, protein and DNA interactions and diurnal and circadian expression patterns are listed in Table 5.2, expression across plant tissues in Table 5.3, and expression responses to different environmental conditions in Table 5.4A. Data on gene expression within different plant tissues and under different conditions was found using the GENEVESTIGATOR online tool (Hruz et al., 2008). Information on circadian and diurnal expression patterns and diurnal phase of gene expression was obtained from the DIURNAL online tool (Mockler et al., 2007). Protein-protein interactions were found using the BioGRID database (Stark et al., 2006). Original sources for the data are referenced in Table 5.4B.

Table 5.2: Functional roles, protein and DNA interactions, and diurnal and circadian expression patterns of transcription factors. Expression pattern data was obtained from the DIURNAL online tool, protein interaction data was obtained from the BioGRID database. Original sources for expression data are listed in Table 5.4B.

			In TF Library		Known Protein Interactions	Known DNA Targets	DNA Motifs Targeted	Expression Pattern				Known Functional Roles In:
	Protein Family	Name	Full CDS	Sequence				Circadian	Phase	Diurnal	Phase	
AT5G61270	**bHLH**	**PIF7**	no	correct	PHYB, PIF3, PIF4, PIL5, TOC1	DREB1C(-)	G-box	-	-	-	-	Far-red light signalling, repression of CBF
AT3G12910	**NAM**	**-**	y	correct	-	-	-	No	-	Yes	4	-
AT3G19290	**bZIP (group A)**	**ABF4**	y	correct	FT, CPK32, CDPK9, CDPK30, CDPK1	-	-	No	-	Yes	0	ABA signalling
AT2G41070	**bZIP (group A)**	**EEL**	y	correct	DPBF2, ABI5	AtEm1(+)	E-box	No	-	Yes	21	ABA signalling
AT3G44460	**bZIP (group A)**	**DPBF2**	y	correct	EEL, ABI5, AREB3, TPR2, At1g10940	-	-	Yes	4	Yes	4	ABA signalling
AT3G23030	**AUX/IAA**	**IAA2**	y	correct	-	-	-	Yes	4	Yes	0	Auxin response
AT4G36730	**bZIP (group G)**	**GBF1**	y	correct	GPRI1, GLK2	-	G-box	Yes	3	Yes	21	Blue light signalling
AT4G34000	**bZIP (group A)**	**ABF3**	y	correct	OST1, GF14 PHI, CPK32	-	-	Yes	0	Yes	2	ABA signalling
AT4G37750	**ERF**	**ANT**	y	correct	-	-	-	Yes	20	Yes	16	Auxin mediated flower development
AT2G21230	**bZIP (group I)**	**AtbZIP30**	y	correct	FT	-	-	No	-	Yes	22	ABA signalling
AT1G06850	**bZIP (group I)**	**AtbZIP52**	y	correct	-	-	-	No	-	Yes	0	-
AT4G38900	**bZIP (group I)**	**AtbZIP29**	y	correct	GRF7, CYCB2;2, CYCB1;3, CDKB1;2, CAK4, KRP2	-	-	No	-	Yes	22	-
AT4G18110	**Zinc finger (C3HC4 RING-type)**	**-**	y	correct	-	-	-	-	-	-	-	-
AT3G61180	**Zinc finger (C3HC4 RING-type)**	**-**	y	correct	-	-	-	No	-	Yes	16	-
AT4G09180	**bHLH**	**FBH2**	y	correct	-	CO(+)	E-box	-	-	-	-	Flowering

Table 5.3: Transcription factor expression in different plant tissues. Absolute expression level categories of low, medium or high were as defined by GENEVESTIGATOR software.

	PIF7	NAM	ABF4	EEL	DPBF2	IAA2	GBF1	ABF3	ANT	AtbZIP30	AtbZIP52	AtbZIP29	(zinc2)	(zinc1)	bHLH
	AT5G61270	AT3G12910	AT3G19290	AT2G41070	AT3G44460	AT3G23030	AT4G36730	AT4G34000	AT4G37750	AT2G21230	AT1G06850	AT4G38900	AT4G18110	AT3G61180	AT4G09180
Protoplast	-	Low	Medium	Low	Low	High	Medium	Medium	Low	Medium	Medium	Medium	-	Medium	-
(root xylem)	-	Low	High	Low	Low	High	Medium	Low	Low	Medium	Low	Medium	-	High	-
(guard cell)	-	Low	Medium	Low	Low	Low	High	High	Low	Medium	Medium	High	-	Medium	-
Conducting Tissue	-	Low	Medium	Medium	Low	High	High	Medium	Low	Medium	Medium	Low	-	Medium	-
Seedling	Low	Low	Medium	Low	Low	High	Medium	Medium	Medium	Medium	Medium	Medium	Low	Medium	Medium
(hypocotyl)	-	Low	Medium	Low	Low	High	High	Medium	High	Medium	Low	Medium	-	Medium	-
(shoot apex)	-	Low	Medium	Low	Low	High	High	High	High	Medium	Low	Medium	-	Medium	-
Roots	-	Low	Medium	Low	Low	Medium	Medium	Medium	Medium	Medium	Medium	Medium	-	Medium	-
Shoot	Medium	Low	Medium	Low	Low	Medium	Medium	Medium	Medium	Medium	Medium	Medium	Low	Medium	Medium
Leaves	Medium	Low	Medium	Low	Low	Medium	Medium	Medium	Low	Medium	Low	Medium	Low	Medium	Medium
(scenescent)	-	Medium	Medium	Low	Low	Medium	Medium	High	Low	Medium	Low	Medium	-	High	-
Inflorescence	Low	Low	Medium	Low	Low	Medium	Medium	Low	Medium	Medium	Medium	Medium	Low	Medium	Medium
(sperm cell)	-	Low	Low	Low	Low	Low	Medium	Medium	Low	Low	Low	Low	-	High	
(ovule)	-	Low	Medium	Low	Low	Low	High	Medium	Medium	Medium	Low	Medium	-	Medium	-
Seed	-	Low	Medium	High	Medium	Low	Medium	Low	Low	Medium	Low	Medium	-	Medium	-
(endosperm)	-	Low	High	High	Medium	Low	Medium	Low	Low	High	Low	Medium	-	High	-

		Hormones		Abiotic Stress		Biotic Stress		Temperature		Light Input	
		Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
AT3G12910	NAM	Ethylene*9, abi1 mutant*6	-	osmotic*7	-	flg22*25*28, oomycete*26, bacteria*27*30*29*33, mpk4 mutant*24	-	-	-	-	-
AT3G19290	ABF4	Zeatin*12, SA*11, ABA*8	GA*10, auxin*2	osmotic*13, salt*15, drought*20*21*22, hypoxia*17*18	-	-	-	heat*12*37	cold*36	far-red*40, red (transient + independent of pif1/3/4/5)*44	-
AT2G41070	EEL	-	-	-	-	-	-	-	-	far-red*40, dark*38*42	-
AT3G44460	DPBF2	-	-	-	-	-	-	-	-	-	-
AT3G23030	IAA2	auxin*2, Zeatin*12	det-2 mutant (BR)*5	salt in roots*13, hypoxia*16*18*19, drought*20*22, nitrate*23	-	mpk4 mutant*24	bacteria*29	-	heat*14*35, cold*34*36	far-red*40, lowR:FR*41, (dark+warm)*38	de-etiolation*45 *46
AT4G36730	GBF1	auxin*2	-	hypoxia*17*18	-	bacteria*29	-	heat*14*37	heat*12, cold*36	far-red*40, dark*42, (dark+warm)*38	light/carbon* 43
AT4G34000	ABF3	ABA*3 *8	auxin*2, Zeatin*12	osmotic*13, salt (roots)*13, drought *21*22, nitrate*23, hypoxia*17	salt*15	-	fungi*33	heat*12	heat*37	light*39, far-red*40, red (transient + independent of pif1/3/4/5)*44	-
AT4G37750	ANT	Zeatin*12	SA*11, ABA*8	-	-	-	-	-	heat*37	-	far-red*40, dark*42
AT2G21230	AtbZIP30	-	MeJa*4	-	-	bacteria*32	-	-	heat*14*37	-	-
AT1G06850	AtbZIP52	-	-	-	anoxia *14, hypoxia*18	-	-	-	heat*14*12	-	de-etiolation*45
AT4G38900	AtbZIP29	auxin*1*2	Zeatin*12	-	-	-	oomycete*31, bacteria*29, fungi*51	-	cold*36	-	-
AT3G61180	-	-	-	hypoxia*17	-	-	fungi*33 (effect via ataf1)	heat*14*37	cold*36	dark*45	-

Table 5.4: Expression responses of transcription factors to different environmental conditions. A: Responses to different conditions and stresses of transcription factors identified in the Y1H screen as able to bind the *LHY* promoter. Numbers preceded by stars are references for original data sources (listed in B). Each colour of text highlights a stress condition of the same type. No information was available for PIF7, FBH2 or AT4G18110. Data was collated using GENEVESTIGATOR software. B: Experimental references for original data in Tables 5.2 and 5.4A.

Table 5.4B

REF NO.	GENEVESTIGATOR REF.	TYPE	ONLINE DATABASE REF.	PUBLISHED
1	AT-00164	Auxin	GSE627	
2	AT-00226	Auxin	GSE3350	Vanneste S et al., 2005
3	AT-00110	Hormones	RIKEN AtGenExpress	Goda et al., 2008
4	AT-00321	Hormones	E-ATMX-13	Pauwels et al., 2008
5	AT-00131	Brassinosteroids	GSE5697	Goda et al., 2008
6	AT-00196	ABA	GSE6151	
7	AT-00199	ABA	E-MEXP-475	
8	AT-00420	ABA	E-MEXP-2378	
9	AT-00113	Ethylene	RIKEN AtGenExpress	
10	AT-00119	GA	GSE5701	
11	AT-00320	SA	GSE14961	
12	AT-00239	Cytokinins	GSE6832	
13	AT-00120	Abiotic Stress	GSE5621	Kilian et al., 2007
14	AT-00230	Abiotic Stress	GSE16222	
15	AT-00403	Abiotic Stress	GSE26983	
16	AT-00201	Hypoxia	GSE14420	Christianson et al., 2009
17	AT-00171	Hypoxia	GSE2218	
18	AT-00447	Hypoxia	GSE21504	
19	AT-00498	Hypoxia	GSE31158	
20	AT-00290	Drought	GSE10643	
21	AT-00292	Drought	GSE10670	
22	AT-00419	Drought	E-MEXP-2377	
23	AT-00266	Nitrate	GSE9148	
24	AT-00414	Defence	E-MEXP-173	
25	AT-00081	Defence	E-MEXP-81	
26	AT-00107	Defence	GSE5615	
27	AT-00128	Defence	E-MEXP-547	Zipfel, 2006
28	AT-00391	Defence	GSE17464	
29	AT-00258	Defence	E-MEXP-739	Michel et al., 2006
30	AT-00169	Defence	GSE2538	
31	AT-00425	Defence	GSE20226	
32	AT-00340	Defence	E-MEXP-1094	
33	AT-00309	Defence	GSE12856	
34	AT-00138	Defence/Temperature	GSE5618	
35	AT-00387	Temperature	GSE18666	
36	AT-00389	Temperature	GSE19254	
37	AT-00402	Temperature	GSE19603	
38	AT-00467	Temperature/Light	E-MTAB-375	
39	AT-00246	Light	GSE7743	Kleine et al., 2007
40	AT-00149	Light	NASCARRAYS-196	Edwards and Millar, 2007
41	AT-00277	Light	GSE9816	
42	AT-00281	Light	GSE10016	
43	AT-00313	Light	E-MEXP-1112	
44	AT-00390	Light	GSE17159	
45	AT-00003	Light	Zimmermann (Gruissem Lab)	
46	AT-00109	Light	GSE5617	

5.2.1.2.1 - Light Signalling Transcription Factors

Two of the transcription factors identified as binding the -957/-754 *LHY* promoter are known to be involved in light signal transduction. GBF1 (G-box Binding Factor 1, AT4G36730) is one of a number of transcription factors mediating cryptochrome (blue-light) dependent photomorphogenic growth in *Arabidopsis* (Mallappa et al., 2006). It is expressed at a moderately high level across all plant tissues (Table 5.3), and displays both diurnal and circadian expression patterns (Table 5.2). Although GBF1 protein is known to be degraded in the dark (Mallappa et al., 2008), GBF1's transcription appears to be up-regulated in the dark and down-regulated by light. These observations fit with its diurnal expression pattern which peaks just before dawn. The timing of GBF1's expression suggests that it may be targeting the *LHY* promoter for activation at dawn in a blue-light-dependent manner.

Also identified as binding the -957/-754 promoter was the transcription factor PIF7 (Phytochrome Interacting Factor 7, AT5G61270). PIF7 is a light-stable regulator that mediates far-red light signalling in *Arabidopsis* through a direct interaction with the Pfr form of PhyB (Leivar et al., 2008). In addition, PIF7 has been recently established as a positive regulator of auxin biosynthesis genes, specifically in the regulation of shade avoidance (Li et al., 2012a), and was also identified as a G-box binding repressor of the CBF pathway genes regulating freezing tolerance (Kidokoro et al., 2009); (Lee & Thomashow, 2012). However, these roles require PIF7 to interact with PhyB, and sequencing results showed that the PIF7 sequence present in the Y1H library was not full-length. The sequence instead corresponded to a known splice variant containing a 5'truncated CDS (Gong et al., 2004). This alternative gene model does not contain the APB motif, which is essential for binding to the Pfr

PhyB protein (Leivar et al., 2008). Therefore, to confirm binding of the PhyB-interacting version of PIF7 to the *LHY* promoter, full-length PIF7 CDS was amplified from cDNA and tested in subsequent yeast interaction assays (Chapter 6). It was later found through these assays that the full-length APB-containing PIF7 did not bind the *LHY* promoter, suggesting that PIF7 does not mediate regulation of *LHY* expression by far-red light. However, this truncation does not affect the ability of PIF7 to interact with TOC1 (Kidokoro et al., 2009). Therefore, the alternative splicing of PIF7 may allow it to have separate roles in both light signalling and the circadian regulation of *LHY*.

5.2.1.2.2 - Abscisic Acid Signalling Transcription Factors

Some of the transcription factors identified as binding the *LHY* promoter are known to have roles in Abscisic Acid (ABA) signalling, these are ABF3 (Abscisic Acid Responsive Element (ABRE) Binding Factor 3, AT4G34000), ABF4 (ABRE Binding Factor 4, AT3G19290), EEL (Enhanced EM Level/DPBF4, AT2G41070) and DPBF2 (ABI5-like 1, AT3G44460) (Kim et al., 2002). The plant hormone Abscisic Acid is important for responses to environmental stresses such as pathogen attack, drought and extreme temperatures, as well as for developmental processes such as seed dormancy and germination (Wang & Irving, 2011).

All four ABA-related transcription factors are diurnally regulated, with expression peaking sequentially around dawn as follows: EEL (ZT-21), ABF4 (ZT-0), ABF3 (ZT-2) and DPBF2 (ZT-4) (Table 5.2). Only ABF3 and DPBF2 have circadian expression, peaking at ZT-0 and ZT-4 respectively. Although neither ABF4 nor EEL show a circadian expression pattern, there may also be some circadian influences on

their expression, since alcohol-induction of the clock gene TOC1 has been shown to down-regulate ABF4 and up-regulate EEL expression (Gendron et al., 2012). Despite these apparent similarities, the responses of these four genes to stress and their expression patterns across plant tissues vary considerably (Table 5.3-5.4), suggesting that these related transcription factors have different functional roles within the plant.

ABF3 and ABF4 are expressed at a much higher level than EEL and DPBF2 in most plant tissues (Table 5.3), and are much more responsive to environmental stresses and plant hormones (Table 5.4A). Expression of ABF3 and ABF4 changes in response to a number of different hormones: ABA, Salicylic Acid (SA), Auxin, Zeatin and Gibberellin (GA). These plant hormones are known to interact both antagonistically and synergistically to regulate diverse processes within the plant (Wang & Irving, 2011). For example, ABA and GA are known to act antagonistically in the regulation of seed germination, with ABA inducing dormancy and GA promoting germination (Gutierrez et al., 2007). ABF4 is both down-regulated by GA and up-regulated by ABA, and shows increased expression levels in the endosperm of seeds (Table 5.3). This could suggest a possible role for ABF4 in the maintenance of seed dormancy or inhibition of germination. However, although ABF3 can delay germination when overexpressed, an equivalent role has not yet been found for ABF4 (Kang et al., 2002).

ABF3 expression is particularly high in the shoot apex and senescent leaves and is down-regulated by both Auxin and the Cytokinin class growth hormone, Zeatin. However, expression of ABF4 is down-regulated by Auxin and up-regulated by

Zeatin (Table 5.4A). Auxins are growth hormones important for cell and organ growth, development and fruiting (Wang & Irving, 2011). In addition, the Cytokinin to Auxin ratio is known to be important for determining cell fate and organogenesis (Moubayidin et al., 2009). The regulation of ABF3 and ABF4 by both hormones therefore suggests a role for these genes in plant development. Auxin-dependent responses including growth and root branching have been reported to be disrupted in 35S-ABF4 plants, indicating that ABF4 may act antagonistically to Auxin-dependent processes (Kang et al., 2002). This correlates with the known antagonism between ABA and Auxin signalling (Wang & Irving, 2011).

ABA is also known to be a key regulator of the plant's drought response, specifically by preventing further water loss through the closure of stomata (Grill & Himmelbach, 1998). ABF3 and ABF4 are known to mediate ABA-dependent responses to drought (Kang et al., 2002). This can also be seen in the up-regulation of ABF3 and ABF4 by ABA treatments, drought and flooding-related stresses, such as altered salt concentrations, changing osmotic pressure and hypoxia (Table 5.4A). ABF3 and ABF4 are known to differentially regulate a wide variety of ABA- and stress-responsive genes, revealing themselves as key mediators of the ABA stress response (Kang et al., 2002). This suggests that targeting of the *LHY* promoter by ABF3 and ABF4 is highly likely to involve the mediation of these ABA-dependent stress responses.

In contrast, EEL and DPBF2 have low expression in most plant tissues but have increased expression levels in seeds. Correspondingly, EEL and DPBF2 are thought to have overlapping roles in seed maturation (Bensmihen et al., 2005), with EEL

acting antagonistically with its homolog ABI5 (ABA-Insensitive-5) in the regulation of genes required for embryogenesis (Bensmihen et al., 2002). It has also been shown that EEL and DPBF2 can interact to form heterodimers *in vitro*, as can EEL and ABI5, and DPBF2 and ABI5. The formation of these heterodimers can alter their individual DNA binding specificity, since the DPBF2-ABI5 heterodimer was shown to be unable to bind an ABI5 target promoter (Kim et al., 2002). It is therefore possible that EEL, DPBF2 and the other DPBF transcription factors, including ABF3 and ABI5, may interact to target the *LHY* promoter in a similar manner. This possibility is explored further in Chapter 6.

5.2.1.2.3 - Auxin

As described in Section 5.2.1.2.2, the transcription factors ABF3 and ABF4 are both negatively regulated by the growth and development hormone Auxin. However, this is not the only route for Auxin signalling to regulate *LHY* expression; IAA2 (Indole-3-Acetic Acid Inducible 2, AT3G23030) was also identified from the Yeast One-Hybrid screen as binding the *LHY* promoter. Most members of the IAA family of transcription factors bind to Auxin Response Factors (ARFs) to repress their activity, and Auxin increases the degradation rates of various IAA proteins to alleviate this repression (Dreher et al., 2006). However, IAA2 is able to bind DNA directly. Any regulation of *LHY* expression by IAA2 would therefore be highly responsive to Auxin.

IAA2 itself is not well characterised, however other IAAs have been suggested to mediate light signalling. Phytochrome A, a photoreceptor for red light, has been reported to interact with and phosphorylate various IAAs including IAA1 *in vitro*

(Colón-Carmona et al., 2000). It has also been shown using Yeast Two-Hybrid that IAA2 and IAA1 can form heterodimers *in vivo* (Kim et al., 1997). IAA2 has both diurnal and circadian expression patterns, peaking at dawn under light-dark cycles and four hours after subjective dawn in constant light. Concordantly, IAA2's expression is up-regulated in the dark and down-regulated on the initial exposure to light of dark-grown seedlings (Table 5.4A). However, it has been suggested that light may act to stabilise IAA proteins (Reed, 2001), which otherwise have extremely short half-lives of around 6-8 minutes (Abel et al., 1994). This suggests that active IAA2 protein is likely to be binding the *LHY* promoter at or just after dawn, and that it could potentially be involved in mediating light signalling to the promoter.

IAA2 expression levels are also affected by hormones. It is induced by both Auxin and Zeatin, opposite to the effect of these hormones on ABF3 (Table 5.4A). This can again be related to the known antagonism between Auxin and ABA signalling. In addition, IAA2 levels are decreased in a *det-2* mutant background (Table 5.4A) deficient in the *DET2* gene required for Brassinosteroid (BR) biosynthesis (Noguchi et al., 1999b, Noguchi et al., 1999a). Regulation of IAA2 by BR is not unexpected, as the Auxin signalling components IAA19 and ARF7 are known to mediate Brassinosteroid regulation of photomorphogenesis (Zhou et al., 2012). This would suggest that IAA2 is up-regulated by BR, and mediates BR-dependent regulation of photomorphogenesis.

Finally, expression of IAA2 is affected by both exposure to pathogens and abiotic stresses, including hypoxia, drought and extremes of temperature (Table 5.4A).

Auxin is well-known to be involved in the plant response to abiotic stresses including cold stress (Shibasaki et al., 2009), oxidative stress and salinity (Iglesias et al., 2010, Jung & Park, 2011). IAA2's response to pathogens varies; its expression is down-regulated by treatment with syringolin, a peptide normally secreted by *P. syringae* (Schellenberg et al., 2010), but is up-regulated in the *mpk4* mutant background, which is associated with elevated SA levels and increased pathogen resistance (Petersen et al., 2000). In addition, Systemic Acquired Resistance (SAR), which provides long-term immunity in tissues adjacent to pathogen exposure, is abolished in various Auxin transport mutants. This effect was suggested to occur through Auxin's impact on SA and Jasmonate (JA) signalling (Truman et al., 2010). Therefore, IAA2 is likely to be mediating both short- and long-term responses to environmental stress, providing another possible route for these signals to feed into the regulation of *LHY*.

5.2.1.2.4 - Flowering-Related Transcription Factors

Some of the transcription factors found to bind the promoter are thought to be involved in the regulation of flowering. These were ANT (Aintegumenta, AT4G37750) and FBH2 (Flowering BHLH 2, AT4G09180).

ANT is an AP2/ERF (Ethylene responsive factors) family transcription factor with known roles in organogenesis, floral development and growth (Elliott et al., 1996, Mizukami & Fischer, 2000). ANT is known to act partially redundantly with another floral regulator, ANT-like 6 (AIL6), and to negatively regulate the expression of the MADS-box regulator Agamous (AG) (Krizek, 2009, Krizek et al., 2000). In accordance with its role in floral development, ANT is highly expressed at the shoot

apex of plants (Table 5.3). ANT is also highly expressed in the earliest stages of organ development, including in floral, cotyledon and leaf primordia (Elliott et al., 1996). ANT is down-regulated by the stress hormones ABA and SA, and up-regulated by the Cytokinin Zeatin, which is primarily involved in growth and development (Table 5.4A). Although ANT expression was not noticeably affected by treatment with synthetic Auxins (Table 5.4A), the effects of the Auxin-inducible gene ARGOS on organ size are dependent on ANT. In addition, treatment with Auxin transport inhibitors decreases the ANT expression level. ANT is therefore thought to act downstream of Auxin in the regulation of organ growth (Hu et al., 2003, Krizek, 2011).

Expression of ANT peaks in the middle of the night under diurnal conditions, at ZT-16. Its expression is also circadian-controlled, peaking at ZT-20 under constant light conditions (Table 5.2). This suggests that any regulation of *LHY* expression by ANT is likely to be repression.

FBH2 was recently identified as one of four novel transcriptional activators targeting the E-box in the *CONSTANS* (*CO*) promoter to promote flowering (Ito et al., 2012). In addition to directly activating *CO*, overexpression of *FBH2* increases levels of the *CO* activators FKF1 (Flavin-Binding Kelch Repeat F-Box 1) and GI, and reduces levels of the negative regulators CDF2 (Cycling DOF Factor 2) and FLC under long days (Ito et al., 2012). Therefore, FBH2 not only directly activates *CO* expression but also modifies the expression of other regulators of *CO*. The circadian clock is also known to regulate *CO*, with ELF3 and LHY down-regulating its expression. This regulation of *CO* by the clock has been directly linked to the timing of

flowering through the photoperiod-dependent activation of *Flowering Locus T* (*FT*) by CO (Suarez-Lopez et al., 2001, Amasino, 2010).

Expression of FBH2 is not rhythmic under constant light conditions or short days. However, it appears rhythmic under long days, with expression peaking around ZT-12, a few hours before dusk. Overexpression of *LHY* is known to repress *CO* expression (Suarez-Lopez et al., 2001). Since FBH2 was found to bind the -957/-754 fragment of the *LHY* promoter (Table 5.1), it can be speculated that FBH2 is likely to repress *LHY* expression, along with other negative regulators of *CO*, to promote the initiation of flowering. However, further investigation is required to confirm this.

5.2.1.2.5 - Pathogen-Defence Transcription Factors

Of the transcription factors identified as binding the -957/-754 *LHY* promoter fragment, only one appeared to be exclusively involved in plant defence against pathogens: AT3G12910. This is a largely uncharacterised NAM (No Apical Meristem) family protein within the NAC superfamily of plant transcription factors, which are involved in hormone signalling, development and biotic and abiotic stress responses (Olsen et al., 2005).

NAM does not appear to be circadian-regulated. However, it has a rhythmic expression pattern under diurnal conditions, with mRNA levels peaking at ZT-4 (Table 5.2). It is expressed at a low but detectable level in most plant tissues, with increased expression in senescent leaves (Table 5.3). Exposure to fungal, oomycete or bacterial pathogens causes up-regulation of AT3G12910 expression (Table 5.4A). It is also up-regulated in the *mpk4* mutant, which exhibits increased pathogen

resistance (Petersen et al., 2000). In addition, it is up-regulated by treatment with Ethylene (Table 5.4A). Ethylene signalling is known to play a role in abiotic stress responses and plant defence, interacting with SA and JA pathways to increase either disease tolerance or programmed cell death in response to pathogens (Wang et al., 2002). This all strongly suggests a role for NAM in the mediation of plant defence responses.

Notably, NAM is co-expressed with two Group III WRKY proteins: WRKY55 (AT2G40740) and WRKY30 (AT5G24110) (as determined using the ATTED-II online tool (Obayashi et al., 2011)). The large and diverse WRKY transcription factor family is involved in many different processes within the plant, including development, plant defence and abiotic stress responses (Eulgem et al., 2000). WRKY30 in particular is a known pathogen response gene in rice species *Oryza sativa* (Peng et al., 2012), and its expression is induced in *Arabidopsis* by reactive oxygen species (ROS), the sudden release of which during the 'oxidative burst' is a key indicator of pathogenic attack (Scarpeci et al., 2008, Bolwell & Wojtaszek, 1997). WRKY proteins are well-known to bind *in vivo* to W-box sequences containing a TGAC core (Eulgem et al., 2000). This sequence is present at three positions in the *LHY* promoter: 867, 584 and 550 basepairs upstream of the translational start site. However, only one of these potential binding sites was present in the Yeast One-Hybrid promoter fragment. Therefore, although these proteins were not detected as binding using the Yeast One-Hybrid screen it is plausible to speculate that they may act with or alongside AT3G12910 to regulate *LHY* expression in response to pathogenic stress. Either way, their association with NAM lends further support to NAM being a part of the plant's pathogen response.

Finally, NAM can be linked to ABA signalling, since although its expression was unaffected by the addition of exogenous ABA, its expression increased in an *abi1* mutant background (Table 5.4A). ABI1 is a negative regulator of ABA signalling (Merlot et al., 2001). Therefore, up-regulation of NAM in the *abi1* mutant suggests that it is at least partially co-regulated with ABA signalling factors, and could potentially act alongside them in mediating stress response signals to the *LHY* promoter.

5.2.1.2.6 - Transcription Factors with Undefined Roles

Three structurally similar bZIP transcription factors from bZIP Group I (Jakoby et al., 2002) were identified as binding the -957/-754 *LHY* promoter in the Yeast One-Hybrid screen: AtbZIP52 (AT1G06850), AtbZIP30 (AT2G21230) and AtbZIP29 (AT4G38900). Little is known of the function of the Group I bZIPs in *Arabidopsis*, apart from one member: VIP1 (AT1G43700, AtbZIP51). A transcriptional activator highly expressed in vascular tissues and siliques, VIP1 was recently shown to have a role in osmotic sensing and is rapidly translocated to the nucleus on submergence of plants in water Tsugama et al. (2012). VIP1 is also known to facilitate infection of the plant by *Agrobacterium*, importing its DNA into the plant nucleus through a direct interaction with bacterial transport (T) complexes (Li et al., 2005).

AtbZIP52 (AT1G06850) is expressed at a relatively low level across the plant, though it is found at moderate levels in flowers, seedlings and roots (Table 5.3). Its expression is not circadian controlled but is rhythmic under diurnal conditions, with a transcriptional peak at dawn. This suggests it is a light-regulated gene, fitting with its down-regulation on exposure to light of etiolated seedlings. Although AtbZIP52

has no known function and does not respond to most hormone or stress treatments, its expression is consistently down-regulated by low oxygen conditions, suggesting a role for AtbZIP52 in abiotic stress responses (Table 5.4A)

AtbZIP30 (AT2G21230) is known to be able to interact directly with the FT protein (Table 5.2), suggesting a role in the regulation of flowering. It is moderately expressed in most plant tissues, with its highest expression level in the endosperm (Table 5.3). Interestingly, this tissue distribution is similar to that of ABF4, which can also directly interact with FT. However, AtbZIP30 does not appear to be responsive to ABA, though it is down-regulated by Methyl Jasmonate (MeJa), a hormone involved in defence, flowering and germination, and is up-regulated after treatment with the bacterial pathogen *P. syringae* (Table 5.4A). Both AtbZIP30 and ABF4 were down-regulated ($p<0.05$) under diurnal conditions by alcohol-induction of TOC1 (Gendron et al., 2012), although neither has a circadian expression pattern. AtbZIP30 does display rhythmic expression under diurnal conditions, peaking just before dawn at ZT-22. The timing of this transcriptional peak suggests that it could potentially be activating *LHY* expression.

AtbZIP29 (AT4G38900) is expressed at a moderate level throughout the plant, though it is particularly highly expressed in the protoplasts of stomatal guard cells (Table 5.3). Its expression is not circadian but is rhythmic under diurnal conditions, peaking at ZT-22, again suggesting that any regulation of *LHY* expression by AtbZIP29 is likely to be activation. Like ABF4, expression of AtbZIP29 is differentially affected by the hormones Auxin and Zeatin (Table 5.4A). However, the direction of this regulation is reversed, with AtbZIP29 being up-regulated by Auxin

and down-regulated by Zeatin. This suggests that these proteins are active under different conditions and may regulate *LHY* expression in response to different signalling pathways. Interestingly, expression of AtbZIP29 appears to have an antagonistic role in plant defence as it is down-regulated by exposure to various pathogens including fungi (*B. graminis*), bacteria (*P. syringae*) and oomycetes (Table 5.4A).

In addition, two zinc finger RING proteins were identified as weakly binding to the -957/-754 fragment of the *LHY* promoter: AT3G61180 and AT4G18110. Almost no data is available for AT4G18110 beyond basic structural predictions, although it is known to have a low level of expression in seedlings, shoots, leaves and flowers (Table 5.3). However, more information is available for AT3G61180, which is expressed at a moderate level across all plant tissues, with particularly high expression in the endosperm, sperm cell and senescent leaves. It has a diurnal expression pattern peaking at ZT-16. Expression is arrhythmic in conditions of constant light and is upregulated in dark-grown seedlings, suggesting that it is negatively regulated by light. The timing of its transcriptional peak suggests a potential role as a repressor of *LHY* expression. It also appears to be regulated by temperature, with its expression being up-regulated by heat and down-regulated by cold conditions. AT3G61180 may be involved in stress responses, since its expression was up-regulated under low oxygen conditions and down-regulated on infection with *B. graminis* (Table 5.4A). Notably, this down-regulation of AT3G61180 by exposure to *B. graminis* was dependent on ATAF1, a NAC family protein known to negatively regulate pathogen resistance that is itself down-

regulated on infection to reduce plant susceptibility (Wang et al., 2009). Therefore, AT3G61180 acts downstream of ATAF1-dependent plant defence signalling.

5.2.1.3 - Functional Analysis of Transcription Factors *in planta*

Whether any of these transcription factors identified as binding the *LHY* promoter had a functional effect on *LHY* transcription was unknown. Therefore, *LHY* transcript levels were determined in SALK mutants of these transcription factors by qPCR of cDNA at dawn (ZT-0) and dusk (ZT-12) as well as subjective dawn (ZT-24) and subjective dusk (ZT-36) (Figure 5.2).

Homozygous SALK lines (NASC) were obtained for nine of the fifteen transcription factors: PIF7, FBH2, DPBF2, EEL, ABF3, GBF1, bZIP52, At3g12910 and At3g61180. SALK lines heterozygous for mutations in the other identified transcription factors were also obtained from NASC. However, these lines are still under selection and have not yet been tested. All tests on these SALK lines were performed by Sally Adams (Carré Lab). The homozygous lines were all found to be null mutants through PCR amplification of transcription factor cDNA.

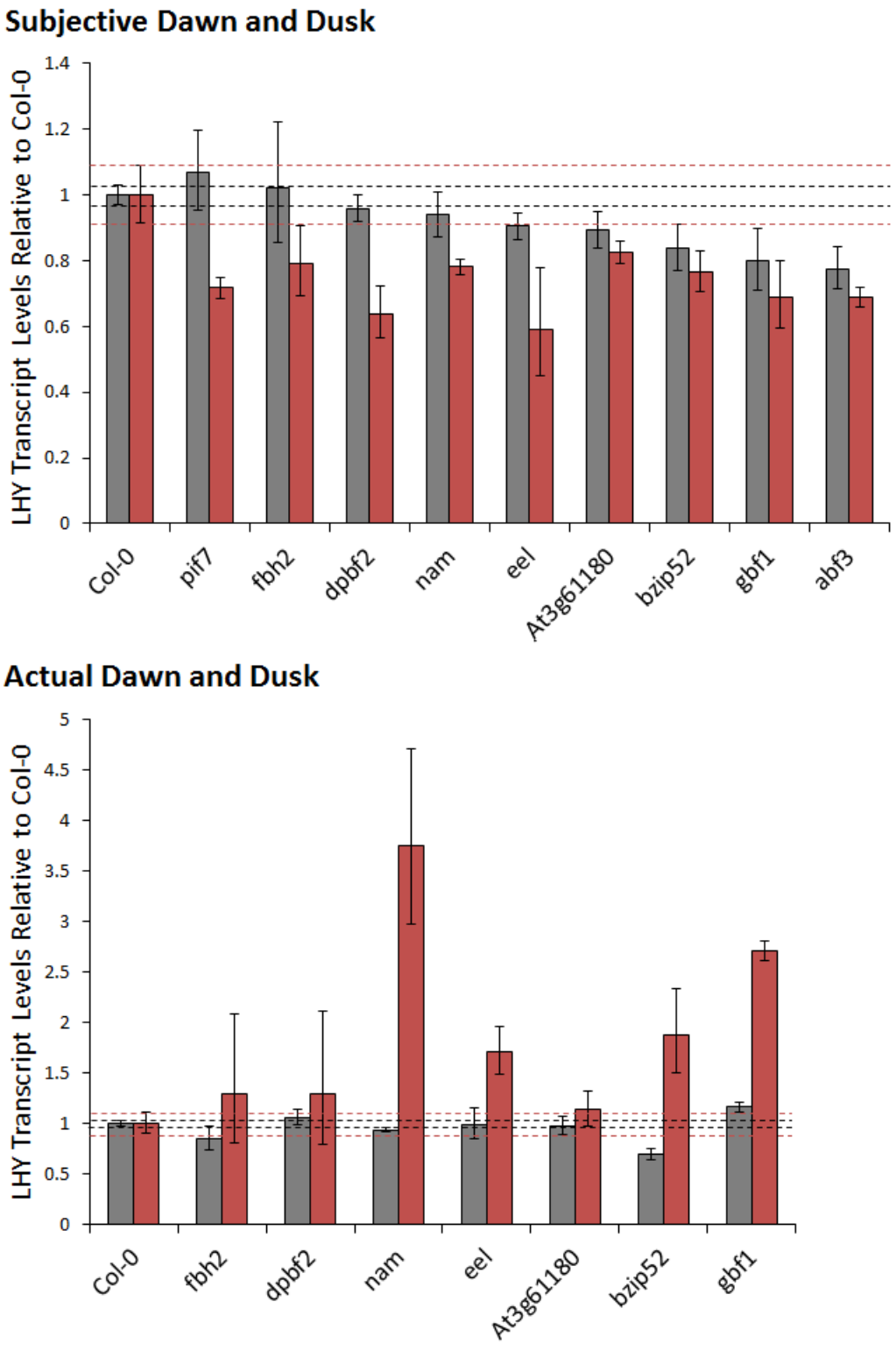


Figure 5.2: *LHY* expression levels are affected in transcription factor knockout plants. *LHY* transcript levels in knockout plants relative to the Col-0 control at subjective dawn (ZT-24, grey), subjective dusk (ZT-36, red), actual dawn (ZT-0, grey) and actual dusk (ZT-12, red). Error bars indicate standard errors. Error range of the Col-0 control is indicated by dashed lines, black for subjective and actual dawn, red for subjective and actual dusk.

The effects on *LHY* transcript levels in these SALK lines at subjective dawn and dusk were minimal compared to the effects seen at actual dusk. The majority of the transcription factors tested here do not display circadian expression patterns, with DPBF2 and GBF1 the exceptions (Table 5.2). The lack of major effect under constant conditions could therefore be caused by rapid dampening of rhythms in constant light. It is also important to note that *LHY* transcript levels at actual dusk are very low. Normalising transcript levels in SALK lines to this control level therefore involves dividing by a very small number, which runs the risk of generating artificially large relative *LHY* levels in response to small increases in actual SALK line *LHY* transcript levels at this timepoint.

All of the SALK lines tested showed slightly reduced *LHY* transcript levels at subjective dusk, as did *bzip52*, *abf3* and *gbf1* at subjective dawn. Since all previously known or expected regulators of *LHY* expression were repressors, these would be the first transcriptional activators to be found to regulate *LHY*. At actual dawn, there were minimal effects again, perhaps due to saturation of activating signals targeting the *LHY* promoter at this time. However, *bzip52* again showed reduced *LHY* transcript levels. At actual dusk, these effects were reversed with *bzip52*, *eel*, *gbf1* and *nam* all showing increased levels of *LHY*. Interestingly, *gbf1* showed increased *LHY* levels at both actual dawn and dusk. This suggests that GBF1 may also function to activate repressors of *LHY*, such as the light-induced *PRR9*. Similarly, it could be speculated that bZIP52, EEL and NAM may also activate day-specific repressors of *LHY*, such that the effect of their absence would only be seen at dusk. However, this is preliminary data and further experiments are required to confirm these conclusions.

5.2.2 - Identifying which Promoter Motifs these Transcription Factors Bind

5.2.2.1 - Yeast One-Hybrid with Mutagenised *LHY* Promoter

The majority of transcription factors identified through Y1H as binding the -957/-754 *LHY* promoter did not have defined DNA binding sites. Therefore, in order to discover which sequence motifs these transcription factors were binding, known and predicted regulatory elements (as described in Chapter 3) within the -957/-754 *LHY* promoter were disrupted through site-directed mutagenesis. These mutated LHY:HISLEU reporter constructs were then individually tested against the Y1H-identified binding factors to discover which sequence motifs were required for binding.

Primers were designed to disrupt individual promoter motifs using the PrimerX online tool (http://www.bioinformatics.org/primerx). The LHY:HISLEU construct was mutagenised using these primers (Methods, Table 2.2), generating multiple *mLHY* promoter constructs with both individual motifs and combinations of motifs mutated (Figure 5.3). The 5A motif and G-box flanking bases mutations (1,2m, 12345m, 345m, St Cl II and Wk Cl II) were generated using previously described primers (Spensley et al., 2009). Primers were also designed to disrupt the core hexamer of the G-box (G-core). In the case of putative Elements 1, 2 and 3, mutations were targeted to the most evolutionarily conserved bases within the motifs (Spensley et al., 2009). Two different mutation sites were generated for both Elements 2 and 3 (ele2m1, ele2m2, ele3m1 and ele3m2) corresponding to their regions of highest sequence conservation. In addition, combinations of different motifs were mutated to test for interactions between adjacent promoter motifs: the G-box with its adjacent $5A_{12}$ motifs, and Element 1 with the $5A_{12}$ motifs.

Each of these mutated -957/-754 *LHY* promoters (*mLHY*) was found to have a different level of auto-activation in yeast. This is likely due to the binding preferences of yeast transcription factors able to activate expression of the HIS reporter gene. Therefore, a negative control (*mLHY* with an empty pDEST22 vector) was included for each promoter to determine the inhibitor level at which false-positives from auto-activation were eliminated. Binding of transcription factors to each *mLHY* promoter strain is shown at its individually determined minimum inhibitor level.

Interestingly, the promoter with all 5A motifs mutated had the lowest level of auto-activation (inhibition by 2.5mM 3AT compared to 25mM for the un-mutated *pLHY* promoter), suggesting that this is a common binding motif for yeast transcription factors. A search of known *Saccharomyces cerevisiae* binding factors (YEASTRACT online database (Abdulrehman et al., 2011, Monteiro et al., 2008, Teixeira et al., 2006)) against these promoter motifs revealed that the binding sites of several yeast transcription factors contain an AAAAA sequence. These included factors known to activate expression of their target genes, such as CUP2p and AZF1p.

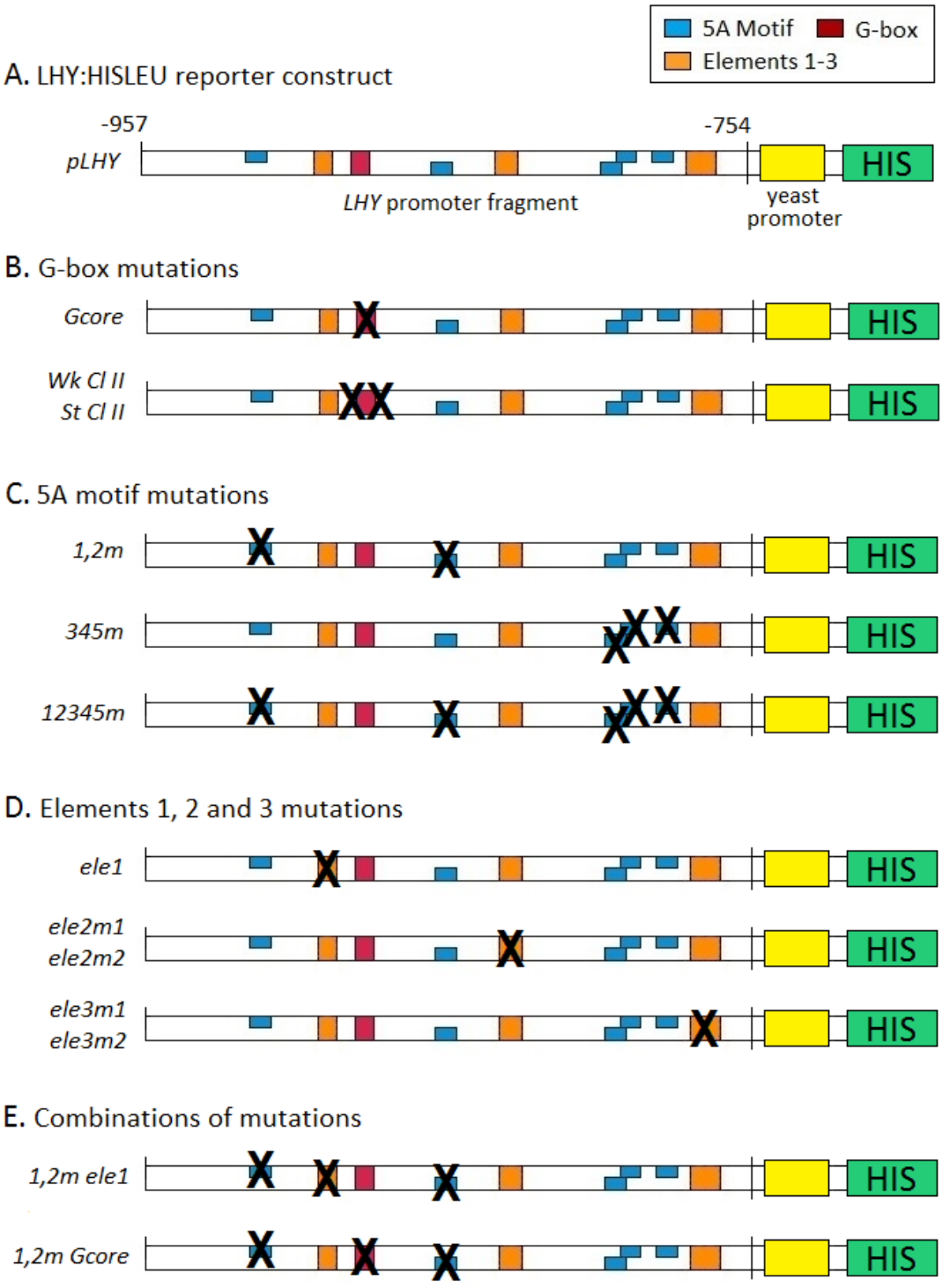


Figure 5.3: Mutated promoter constructs for Yeast One-Hybrid. LHY:HISLEU reporter constructs, containing the -957/-754 region of the *LHY* promoter with various conserved elements disrupted were generated by site-directed mutagenesis. The basepair substitutions for each mutated construct are detailed in Table 2.2. X denotes a disrupted promoter element. Italics indicate construct names as referred to in Table 2.2 and Figure 5.4.

5.2.2.2 - Transcription Factors Bind Specific Motifs on the *LHY* Promoter

Most of the transcription factors tested displayed interactions that were too weak to accurately determine their target motifs in this assay. However, six transcription factors gave clear results: GBF1, ABF3, ANT, NAM, bZIP29 and bZIP30 (Figure 5.4).

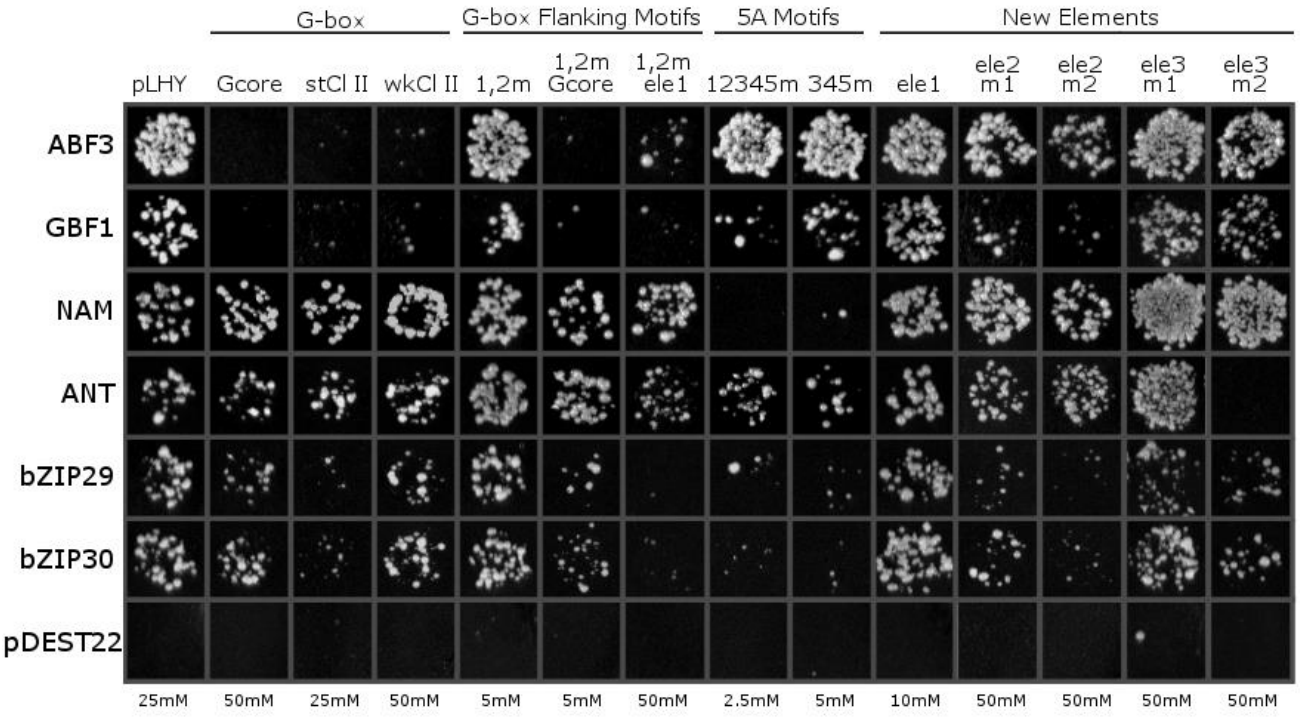


Figure 5.4: Binding affinity of transcription factors at the -957/-754 *LHY* promoter. Haploid Y187 yeast containing mutated LHY:HISLEU reporter constructs were independently mated with AH109 yeast containing transcription factors in pDEST22 vectors. Auto-activation levels were determined separately for each mutated promoter as the inhibitor concentration (mM 3AT below each promoter) at which growth was abolished for a mLHY:HISLEU + empty pDEST22 vector control. Promoter mutations are shown in Figure 5.3.

Binding of ANT to the promoter was abolished by Element 3 mutation 2, which disrupted the sequence CTTC. This is the first functional role found for Element 3, which had been identified only as a region of sequence conservation across species (Spensley et al., 2009). ANT's DNA binding specificity was previously defined as the sequence 5'-GCAC ($^{G}/_{A}$) N ($^{A}/_{T}$) TCCC ($^{G}/_{A}$) ANG ($^{T}/_{C}$)-3' (Nole-Wilson & Krizek, 2000), which contains the CTTC motif as the complementary sequence 5'-($^{G}/_{A}$) CNT ($^{T}/_{C}$) GGGA ($^{T}/_{A}$) N ($^{C}/_{T}$) GTGC-3'.

The NAM-family protein AT3G12910 had no known DNA binding site. Through this mutagenesis Yeast One-Hybrid assay, its binding to the *LHY* promoter was found to be dependent on a single type of sequence motif: the 5A motif. NAM was unable to bind when all 5A motifs were disrupted, and its binding was severely weakened when the proximal 5A ($5A_{345}$) motifs were mutated. However, mutation of the distal 5A ($5A_{12}$) motifs did not affect binding to the promoter, suggesting that NAM binds the *LHY* promoter through the three proximal 5A motifs.

ABF3's binding to the *LHY* promoter required the G-box to be intact. This was expected as the ABFs are known to target G-box containing sequences (Choi et al., 2000). Its binding was also weakened by the combined mutation of the three motifs immediately adjacent to G-box: Element 1 and the two distal 5A motifs ($5A_{12}$). Since ABF3's binding was unaffected by single mutation of either $5A_{12}$ or Element 1, ABF3's interaction with the G-box is not dependent on either of these motifs. Rather it suggests that ABF3 can form a secondary interaction with Element 1 and the $5A_{12}$ motifs, and that the presence of either one of these flanking motifs is sufficient to stabilise its binding at the G-box.

GBF1 is also known to bind to motifs containing an ACGT sequence, including G-boxes (Schindler et al., 1992). In this assay, binding of GBF1 to the *LHY* promoter was found to require the G-box. Like the G-box-binding ABF3, GBF1 also appears to interact with one of the adjacent Element 1 and $5A_{12}$ motifs, most likely to stabilise its binding at the G-box. Unlike ABF3, mutation of all five 5A motifs was sufficient to weaken GBF1's binding, suggesting that the 5A motifs are stabilising GBF1 binding at the G-box. Binding of GBF1 was also disrupted by Element 2 mutation 2, which destroyed the A-rich sequence AAATAAA. This suggests that binding of GBF1 is mediated by the G-box and by A-rich elements.

Two of the transcription factors had binding preferences that were less clear-cut, these were bZIP29 and bZIP30. Their binding was unaffected by mutation of the G-box core hexamer, but was weakened by mutation of flanking sequences (St Cl II). Like the G-box binding proteins ABF3 and GBF1, both bZIP29 and bZIP30 appeared unable to bind in the combined absence of the Element 1 and $5A_{12}$ motifs, but were unaffected by these mutations alone. Interestingly their binding was not disrupted by the Wk Cl II mutation, which had alterations to the same flanking bases as St Cl II (G-box: ACCACGTGTC, St Cl II: GTCACGTGAC, Wk Cl II: CTCACGTGAG). This suggests that the identity of these G-box flanking nucleotides is therefore important for the binding of these bZIP transcription factors.

Despite their similarities, the binding specificities of bZIP29 and bZIP30 are not identical. Binding of bZIP29 was weakened by disruption of the GAGA sequence within Element 3 (ele3m1 mutation), whereas binding of bZIP30 was unaffected. In addition, although binding of both was significantly weakened by the loss of

Element 2, bZIP29 required both CCTC and AATA sequences within Element 2 to be intact (i.e. no binding was detected with either ele2 mutation) whereas bZIP30 only required AATA (i.e. it was able to bind despite the ele2m1 mutation). It is important to note that this AATA sequence within Element 2 is part of a longer sequence bearing some similarity to the 5A motifs: AAATAAA. In addition, mutation of all five 5A motifs in the promoter reduced binding of both bZIP29 and 30. This suggests that an A-rich sequence is required for bZIP30 and bZIP29 binding.

5.3 - Discussion

5.3.1 – Summary of Y1H Results

We have identified through Yeast One-Hybrid a number of transcription factors capable of binding to the -957/-754 *LHY* promoter. These include proteins with known roles in a broad range of plant processes, including light signalling, flowering, growth, abiotic stress responses and plant defence (Tables 5.2-5.4). The identification of these transcription factors provides a novel route for wide-ranging environmental and developmental signals to be integrated into the circadian clock (Figure 5.5). In addition, many of these transcription factors appear to be activators of *LHY* expression (Section 5.2.1.3). These are the first regulators to be found that can activate *LHY* expression, since all previously identified regulators were repressors.

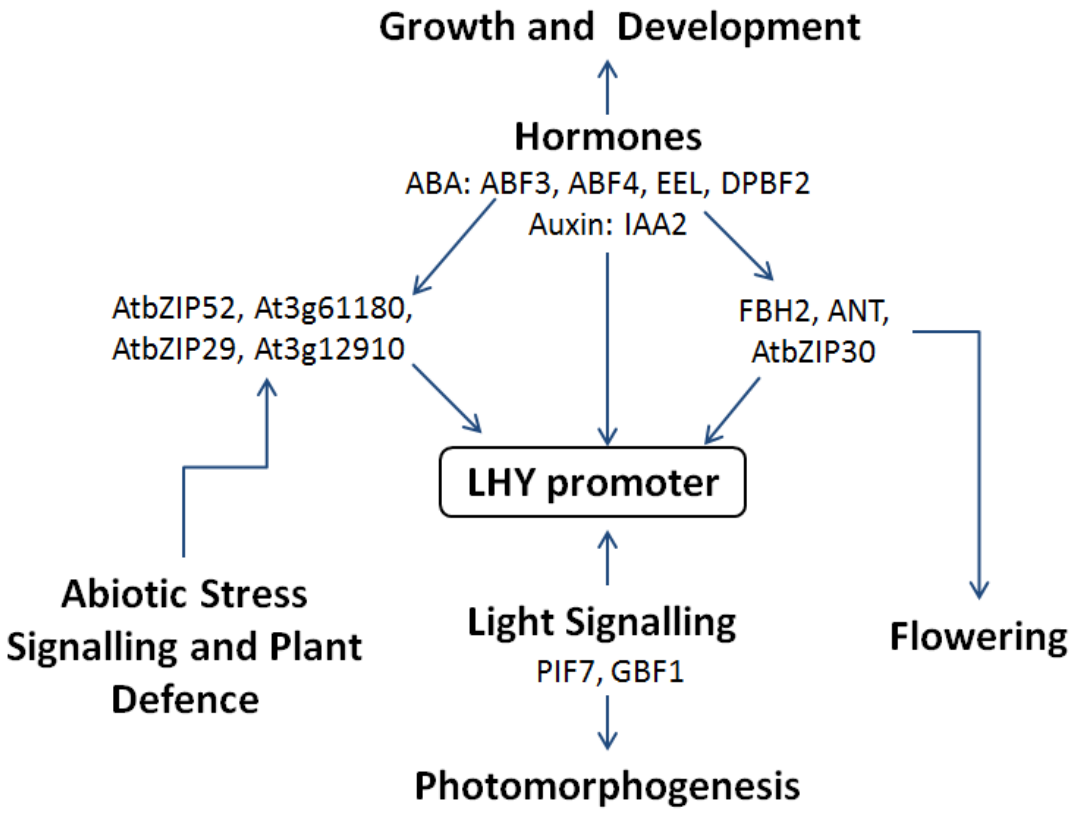


Figure 5.5: Pathways feeding into the *LHY* promoter. Transcription factors targeting the -957/-754 region of the *LHY* promoter are involved in a wide variety of biological processes within the plant, and enable hormone, light and stress signals to feed into the clock through *LHY*.

We have also been able to identify the binding sites of six of these transcription factors (Figure 5.6), and in doing so have assigned the first functional roles to the evolutionarily conserved Elements 1-3. We have also been able to provide further evidence for interactions between different promoter motifs, which was initially suggested by data from *luciferase* assays (Chapter 3). Most notably, the 5A motifs appear to be involved in stabilising the interaction of GBF1 and ABF3, and possibly of bZIP29 and bZIP30, with the G-box.

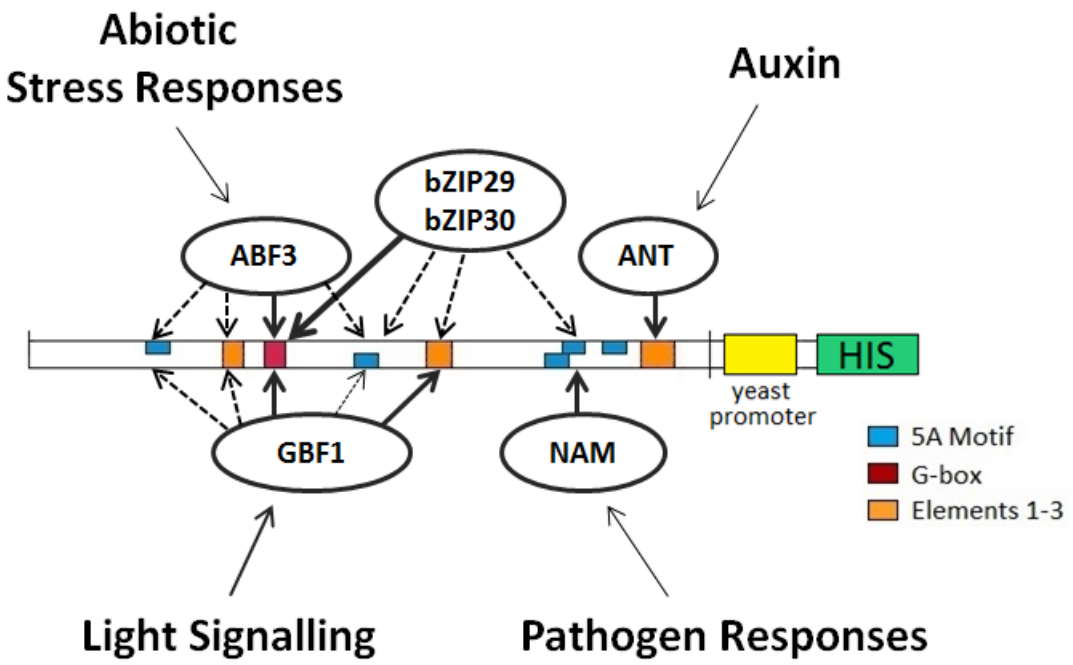


Figure 5.6: Environmental and developmental signals feed into the *LHY* promoter through transcription factors binding to specific promoter motifs. Solid arrows indicate that a motif is required for binding of a transcription factor. Dashed arrows indicate weaker stabilising interactions with motifs.

5.3.2 - Light Signalling Transcription Factors Can Bind the *LHY* Promoter

All of the identified transcription factors for which expression data is available have rhythmic expression patterns under diurnal conditions, suggesting that they are light regulated. In addition, two transcription factors (PIF7 and GBF1) are known to play functional roles in red and blue light signalling pathways, providing a route for regulation of *LHY* expression by light. Interestingly, the role of PIF7 is complicated by the different binding specificities of its splice variants. The full-length PIF7 (gene model 1) which is able to directly interact with the far-red light photoreceptor PHYB, was not able to bind to the *LHY* promoter in these assays. However, truncated PIF7 (gene model 2) which does not contain the PHYB binding site, was able to bind strongly to the *LHY* promoter. Alternative splicing is known to play an important role in clock-regulated temperature compensation, with reversible

alternative splicing of clock genes including LHY, CCA1, PRR9, PRR7, PRR5 and TOC1 occurring in response to temperature shifts. It is thought that the accumulation of transcript splice variants unable to form functional protein effectively enables the transient down-regulation of these genes without having to alter broader transcriptional processes (James et al., 2012). Whether this alternative splicing of PIF7 is temperature-dependent and how it affects the regulation of the clock requires further investigation.

5.3.3 - Developmental Signals Feed Into the *LHY* Promoter

Expression of many of the transcription factors responds to hormones such as Auxin, GA, BR and Zeatin, which are involved in the regulation of photomorphogenesis, growth and development (ABF4, ABF3, GBF1, bZIP29, IAA2 and ANT). In addition, several can be directly linked to flowering, through either the development of floral organs (ANT), the regulation of CO (FBH2), or direct interaction with FT (bZIP30 and ABF4). The flowering regulator FLC had been previously identified as binding to the *LHY* promoter *in planta*, a result that is replicated here, although its inconsistent binding in the yeast assays suggests that stable binding may require the presence of co-factors. The binding of these transcription factors to the *LHY* promoter therefore provides further evidence that signals involved in the timing and initiation of flowering feed into the clock through the regulation of *LHY* expression.

5.3.4 - Environmental Stress Signals Feed Into the *LHY* Promoter

The majority of the transcription factors identified as binding the *LHY* promoter were responsive to environmental stresses, in particular drought-related abiotic stresses such as hypoxia, osmotic stress and salt stress (ABF4, ABF3, GBF1, bZIP52, IAA2,

NAM and At3G61180). In addition, many of the transcription factors (most notably NAM, as well as IAA2, bZIP30, bZIP29, At3G61180, GBF1 and ABF3) had altered expression in response to various fungal and bacterial pathogens. In keeping with these roles, several of these transcription factors were responsive to plant hormones involved in defence and stress response signalling such as ABA, SA and JA (ABF3, ABF4, bZIP30 and ANT). Temperature also appears to be important for the regulation of most of the transcription factors. It is therefore clear that abiotic and biotic stress signals can feed into the clock through the regulation of these *LHY* promoter-binding transcription factors.

Abscisic acid signalling components (ABF4, ABF3, EEL, DPBF2) and responsive genes (ABF4, ABF3, ANT) were particularly well-represented in the collection of *LHY* promoter-binding transcription factors. In addition, the defence-related gene *NAM* was up-regulated in an *abi1* mutant (Table 5.4A). Since ABI1 is a negative regulator of ABA signalling, this suggested that *NAM* may be co-regulated with ABA-related genes. This over-representation of ABA-related proteins at the *LHY* promoter therefore suggests that ABA-mediated biotic stress signalling feeds into the clock through regulation of *LHY*.

5.3.5 - Regulatory Feedback Between the Clock and Proteins Targeting the *LHY* Promoter

While these transcription factors can regulate *LHY* expression, there is also evidence for regulatory feedback from the circadian clock. Several of the transcription factors have circadian rhythms of expression in constant light (ABF3, DPBF2, GBF1, IAA2 and ANT), and others have altered expression on alcohol-induction of TOC1

(*bZIP30* and *ABF4* being down-regulated and *EEL* being up-regulated by TOC1). In addition, a positive feedback loop can be identified between the clock and ABA signalling: the ABA signalling factors ABF3, ABF4, EEL and DPBF2 were found to bind the *LHY* promoter, and mRNA levels of *LHY* were reduced in *abf3*, *eel* and *dpbf2* mutants, suggesting that they activate *LHY* transcription. In turn, *ABF3* expression in *LHY* over-expressing plants was up-regulated at its transcriptional peak (DIURNAL online tool (Mockler et al., 2007)), suggesting that it is regulated positively by LHY.

This type of positive feedback loop would be particularly useful in the amplification of stress signals, i.e. ABA is induced in response to environmental stress and up-regulates ABF3 expression, ABF3 could then up-regulate LHY, which increases the levels of ABF3 within a limited time-frame. This enables both ABF3 and LHY levels to remain elevated for an extended but finite period of time. This would suggest that the clock mediates acute and transient amplification of stress signals through LHY, enabling them to accumulate sufficiently to alter the expression of appropriate downstream stress response genes.

The presence of a positive feedback loop between ABA-mediated stress responses and LHY adds credence to the theory of feedback from the clock to its other input signals, such as light (GBF1), Auxin (IAA2) and flowering and development (ANT). It is important to note that many of the transcription factors identified as binding the *LHY* promoter have functions in more than one signalling pathway. Cross-talk between different hormone signalling pathways is well known to occur in plants (Wang & Irving, 2011). It is therefore probable that at least some of these

transcription factors may interact at the promoter to regulate *LHY* expression. This possibility is explored further in Chapter 6.

CHAPTER 6

Transcription Factor Interactions at the *LHY* Promoter

6.1 - Introduction

As described in Chapter 5, a Yeast One-Hybrid screen performed on the -957/-754 *LHY* promoter fragment identified fifteen transcription factors as capable of binding the promoter. In addition, the binding of six of these was mapped to one or more specific promoter motifs (Figure 5.4). Many of these transcription factors were involved in pathogen and abiotic stress responses, with expression of several genes responsive to the stress hormone ABA. As detailed in Chapter 5, some of these transcription factors were known to be capable of protein-protein interactions with one another, including several of the ABA-related proteins. Other transcription factors were known to be involved in interactions with clock proteins, such as PIF7's ability to bind TOC1 protein (Kidokoro et al., 2009). It is therefore likely that these identified binders of the *LHY* promoter interact both with each other and with other transcription factors targeting the promoter.

Regulatory interactions between different regions of the *LHY* promoter had been inferred from *in planta* luciferase analysis (Chapter 3). In particular, the proximal 5A motifs were proposed to relieve repression in the G-box-containing distal promoter region. The G-box was also identified as mediating both direct and indirect regulatory effects of TOC1 on *LHY* expression (Chapter 4). These results indicated the presence of multiple regulatory interactions between transcription factors binding the *LHY* promoter. These could include both synergistic interactions, whereby one transcription factor aids the binding or function of another at the promoter, and

antagonistic interactions, where the binding of one transcription factor is prevented by others.

Clock transcription factors were the most obvious candidates for involvement in such regulatory interactions at the promoter. However, the broad functional roles of the 15 identified binding factors, from pathogen responses to light signals, developmental cues and environmental stresses, suggested there was potentially a much wider network of transcription factor interactions centring on the *LHY* promoter. Therefore, to elucidate the details of this regulatory network, a range of transcription factors was selected for an investigation of protein-protein interactions at the *LHY* promoter. This selection comprised some of those transcription factors known or suspected to bind the promoter, as well as other functionally related transcription factors.

Due to the considerable set-up time required for this interactions assay, termed a modified-Y1H assay (mY1H), it was not possible to test all the transcription factors identified from the Y1H screen. Therefore, only the strongest and most consistent binders across both mating and transformation based assays (Table 5.1) were selected. These were ABF3, NAM and PIF7. However, when the PIF7 CDS from the Y1H collection was sequenced, it was found to contain a 5' truncation of its CDS. This truncated CDS corresponded to a known and relatively unstudied splice variant of PIF7, lacking the phytochrome-interacting domain. Nevertheless, PIF7 was still a likely candidate for targeting the *LHY* promoter, both as part of the light signalling input pathway and potentially also the mediation of TOC1 binding, since both full-length and truncated PIF7 can bind TOC1 (Kidokoro et al., 2009). PIF7 was

therefore re-amplified from *Arabidopsis* cDNA and included in the modified-Y1H assay as full-length CDS. The light signalling transcription factor GBF1 was also identified as a strong binder of the *LHY* promoter from yeast assays. However, it was not included in the mY1H assay as PIF7 was considered a more potentially interesting light signalling candidate due to its interaction with TOC1.

Several ABA signalling pathway proteins had been identified as binding the *LHY* promoter from the Y1H screen, including EEL, DPBF2, ABF4 and ABF3. Although only ABF3 bound strongly and consistently enough to be tested further in the modified-Y1H assay, the influence of ABA on promoter regulation was highly likely to involve other transcription factors. The ABA-responsive transcription factor ABI5 is known to have some overlapping functions with ABF3, particularly in stress responses of seedlings (Finkelstein et al., 2005). In addition, another ABA-related protein, ABI3, was already known to physically interact with both ABI5 and TOC1 in yeast (Nakamura et al., 2001, Kurup et al., 2000). Both ABI3 and ABI5 were therefore speculatively included in the mY1H assay.

The flowering-related MADS-box protein FLC had previously been shown to weakly bind to the *LHY* promoter *in planta* through ChIP. This was suspected to occur through the 5A motifs since they resembled the MADS-box binding site (Spensley et al., 2009). However, FLC appeared to have difficulty binding the promoter in yeast, with binding apparent in only one yeast assay. Such inconsistent binding could be due to the absence of appropriate binding co-factors in yeast. Alternatively, it may indicate that FLC is not the primary MADS-box transcription factor to target the 5A motifs of the *LHY* promoter. Another transcription factor from

the MADS-box family (MADS44) was identified as a very strong binding target for LHY from ChIP-seq data (Adams, Veflingstad and Carré, unpublished). This suggested the possibility of a transcriptional feedback loop between MADS44 and LHY. Both MADS44 and FLC were therefore included in the mY1H assay to test for transcription factor interactions at the *LHY* promoter.

The circadian clock transcription factors PRR9, PRR7 and TOC1 had all at this time been shown to bind the *LHY* promoter through ChIP (Chapter 4) (Nakamichi et al., 2010). However, they were not found to bind the -957/-754 region of the *LHY* promoter in either the mating- or transformation-based yeast assays, suggesting that their binding may require co-factors. Since PRR-family proteins are known to be capable of interacting with one another, PRR9, PRR7 and TOC1 were included in the modified-Y1H assay to investigate any potential protein interactions at the promoter.

LHY itself was not thought to act on its own promoter at this time, and had not been identified as binding in earlier yeast assays. However, subsequent to these initial tests, ChIP-seq data showed that LHY binds its own promoter *in planta* (Adams, Veflingstad and Carré, unpublished). LHY was therefore selected as the final transcription factor for the modified-Y1H assay to investigate whether its binding might require co-factors.

Eleven transcription factors were therefore selected for testing against one another in the modified-Y1H assay, encompassing three proteins from the original Y1H screen (ABF3, NAM and PIF7), a flowering-related protein (FLC) previously shown to

bind the *LHY* promoter in ChIP, a MADS-box transcription factor known to be a binding target of LHY (MADS44), four clock-related proteins shown to bind in ChIP (PRR9, PRR7, TOC1) or ChIP-seq (LHY), an ABA signalling protein known to interact with TOC1 (ABI3) and an ABA-responsive transcription factor known to interact with ABI3 (ABI5).

Aims

- Establish a modified-Y1H assay as a method for investigating interactions of specific transcription factors at the *LHY* promoter (Section 6.2.1)
- Identify antagonistic and synergistic interactions between specific transcription factors at the *LHY* promoter using a modified-Y1H assay (Section 6.2.2).
- Investigate how these interactions may impact on the regulation of *LHY* expression *in planta* (Section 6.2.3)

6.2 - Results

6.2.1 - A Modified Yeast One-Hybrid Assay (mY1H) as a Method of Investigating Protein-Protein Interactions at the *LHY* Promoter

A modified Yeast One-Hybrid assay (modified-Y1H, mY1H) was performed to investigate antagonistic and synergistic protein interactions at the *LHY* promoter. This involved performing standard Yeast One-Hybrid assays (testing the binding of one protein of interest to one promoter through activation of a reporter gene) in cells that were also transformed with a second Transcription Factor (TF) lacking a transcriptional activation domain and therefore unable to activate the reporter gene. Comparisons to controls without this additional TF allow us to draw conclusions

about the interactions between the two proteins at the promoter, i.e. whether the second transcription factor assists or prevents binding of the first transcription factor to the promoter (Figure 6.1).

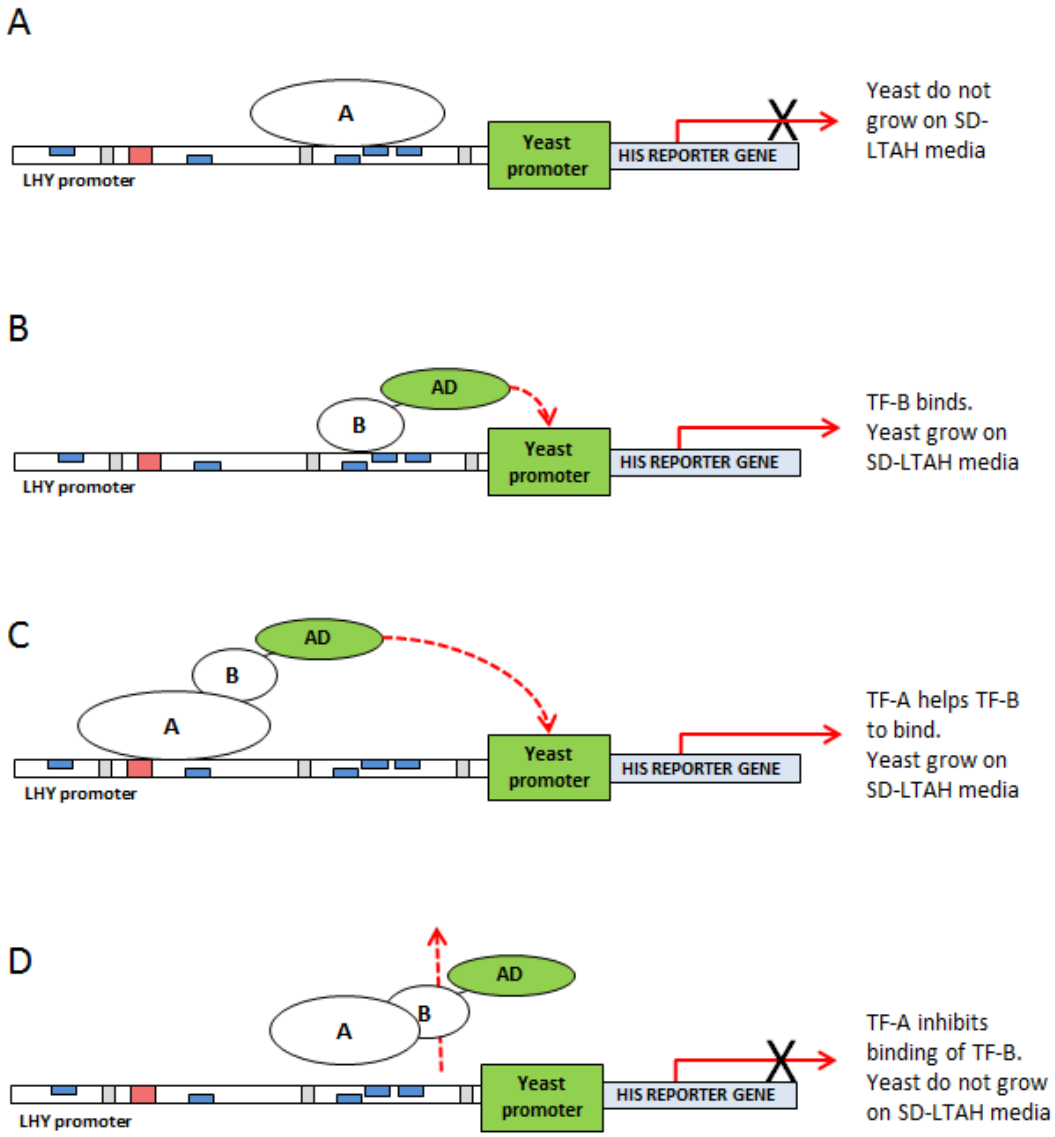


Figure 6.1: The modified Yeast One-Hybrid assay. Two transcription factors (A and B) were transformed into a diploid pLHYGold *–LTAH* yeast strain, containing the -957/-754 *LHY* promoter fragment in a HIS reporter construct. A: Transcription factors lacking a yeast activation domain (TF-A) are unable to initiate transcription of the reporter gene. B: Transcription factors with yeast activation domain (TF-B) can initiate transcription of the reporter gene when bound to the promoter. C: TF-A mediates binding of TF-B, allowing it to activate reporter gene. D: TF-A inhibits binding of TF-B, preventing activation of the reporter gene.

6.2.1.1 - Methodology

Transcription factors were cloned into yeast vectors with and without the yeast Activation Domain (AD), which is required for activation of the reporter gene. Vectors pDEST22 (*trp* selection) and pARC352 (*ade* selection) were used for transcription factors with (TF+AD) and without (TF-AD) yeast activation domains respectively. These TF +/- AD constructs were sequentially transformed in all possible combinations into a yeast strain containing pLHY:HISLEU, such that each yeast strain contained the -957/-754 *LHY* promoter (*leu* selection), one TF+AD (*trp* selection) and one TF-AD (*ade* selection). Since the AD is required for activation of the *HIS* selective marker gene, only binding of TF+AD to the promoter could allow yeast growth on *-leu-trp-ade-his* (SD-LTAH) selective media. Therefore, any change in growth must be due to the effect of TF-AD on the binding of TF+AD.

In this assay, stabilisation of TF+AD binding by TF-AD or recruitment to the promoter of TF+AD by TF-AD would result in increased yeast growth on SD-LTAH selective media. Conversely, inhibition of TF+AD binding by TF-AD, through direct competition for a binding site, blocking of a binding site by TF-AD, or other antagonistic protein-protein interactions, would cause inhibition of yeast growth on SD-LTAH selective media. GFP was also cloned into these vectors and used as a negative control against transcription factors in both vectors. Any errors found in transcription factor sequences were corrected through re-amplification of coding sequences from cDNA or site-directed mutagenesis prior to the assay.

6.2.1.2 - Optimisation of Assay Conditions

To enable comparison of TF-promoter interactions across different mY1H yeast strains, a quantitative approach was taken to the modified-Y1H assay. Growth rates of *LHY* promoter strains containing different transcription factors were established over a range of auto-activation inhibitor (3AT) concentrations. In order to visualise the interactions, the tests were also carried out at different cell concentrations.

Initial growth tests were conducted with yeast strains containing the *LHY* promoter and one of the following transcription factors from the Y1H screen results: ABF3, ABF4, EEL and DPBF2. Two independent cultures of each were grown on SD-LT agar and tested at the following range of plated cell concentrations: 10^8, 10^7, 10^6, 10^5 and 10^4 cells per volume plated (3µl). As in the previous yeast-based assays, 3AT inhibitors were added to assess growth and strength of binding in the absence of auto-activation of the *LHY* promoter. Inhibitor concentrations of 5, 25, 50, 75 and 100mM 3AT were tested on SD-LTH plates.

Integrated intensity measurements extracted from photographs using MetaMorph software were used to plot growth of each strain after 3 days (Figure 6.2). No difference in growth was seen between different yeast strains on either SD-LT or SD-LTH plates, indicating that amount of growth can be compared across yeast strains containing a different transcription factor.

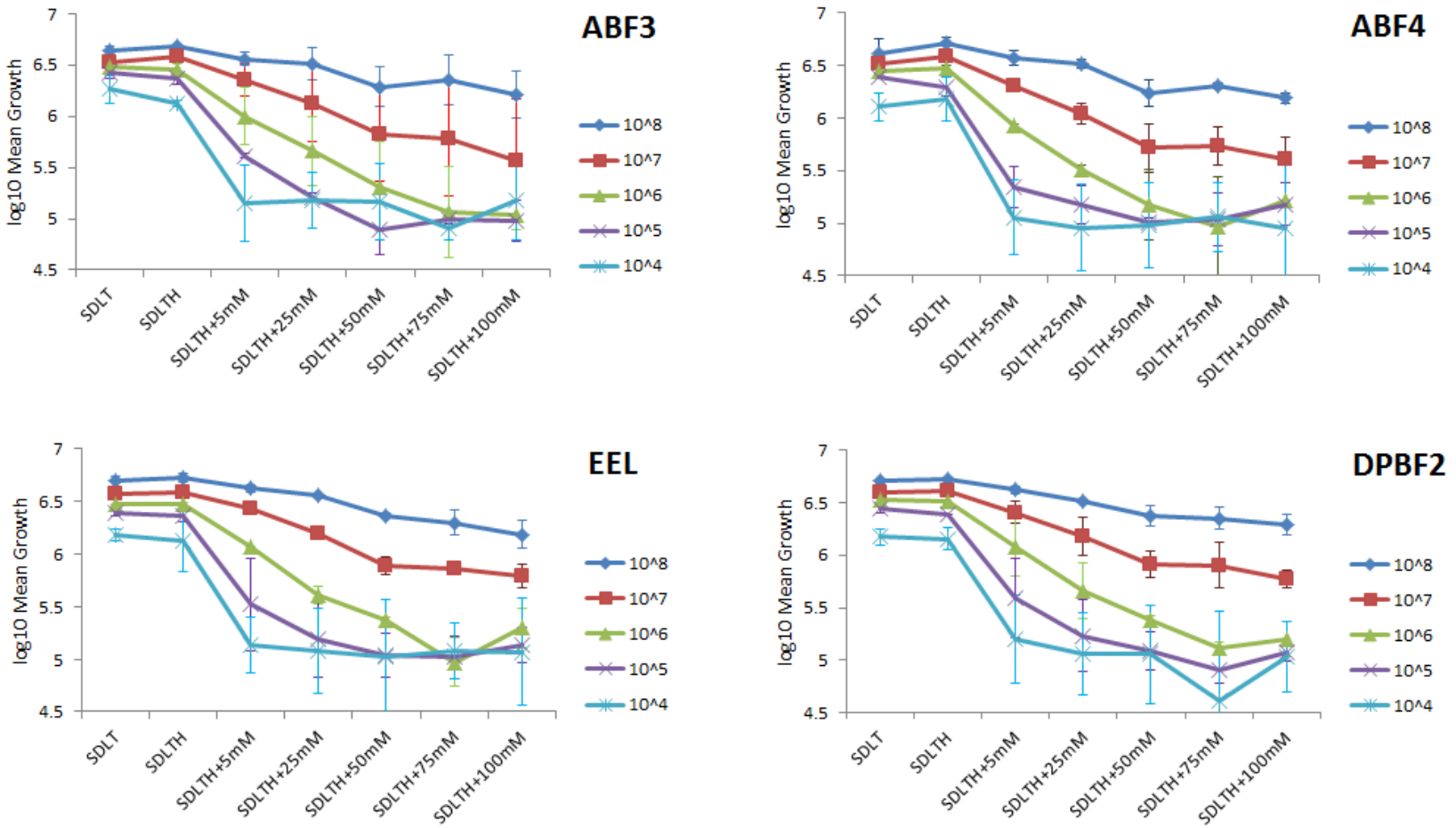


Figure 6.2: Growth rates are constant across yeast strains. LHY:HISLEU constructs were transformed into diploid yeast with either ABF3, ABF4, EEL or DPBF2 in pDEST22 (*trp* selection) vectors. Each strain was grown from a starting amount of 10^8-10^4 cells, and plated on SD-LTH media containing between 0-100mM 3AT inhibitors. Integrated intensity measurements of the growth of each yeast strain after 3 days were extracted from desaturated photographs using MetaMorph software and background corrected to a region with no growth. Data shown are log10 means across two biological replicates. Error bars are standard deviations.

Yeast plated at the 10^8 cell concentration showed largely uniform growth across all the plates, making this concentration unsuitable for use in the modified-Y1H assay. Conversely, yeast plated at 10^4 and 10^5 cell concentrations showed dramatic reductions in growth at just 5mM inhibitors, indicating that they were also unsuitable for the modified-Y1H assay. 10^6 and 10^7 cells plated showed the gentlest gradation from high to low growth across the inhibitor range, with both concentrations showing a steady decrease in growth from 0-50mM inhibitors. Therefore, 10^6 and 10^7 cells on inhibitor concentrations of 0, 5, 25 and 50mM were selected as the optimal conditions for the modified-Y1H assay, since this growth range should allow detection of both increased and decreased growth due to transcription factor interactions.

6.2.1.3 - Description of mY1H Figures

Figure 6.3 A-H shows the growth of multiple different yeast strains, each containing the *LHY* promoter and two transcription factors, on selective plates. Each figure shows the individual effects of each transcription factor (without activation domain, TF-AD) on the binding of a single transcription factor (with activation domain, TF+AD) to the *LHY* promoter. As in previous yeast assays, growth under *–his* (H) conditions indicates an interaction between promoter and TF+AD, since without the yeast activation domain the transcription factor (TF-AD) is unable to activate expression of the HIS reporter gene. The figures display technical replicates of each strain plated at two different cell concentrations (10^6 or 10^7 cells per 3µl plated) from a single starting culture. These cell concentrations were established as the most appropriate for observing growth effects as described in Section 6.2.1.2.

A: NAM strengthens binding of ABF3

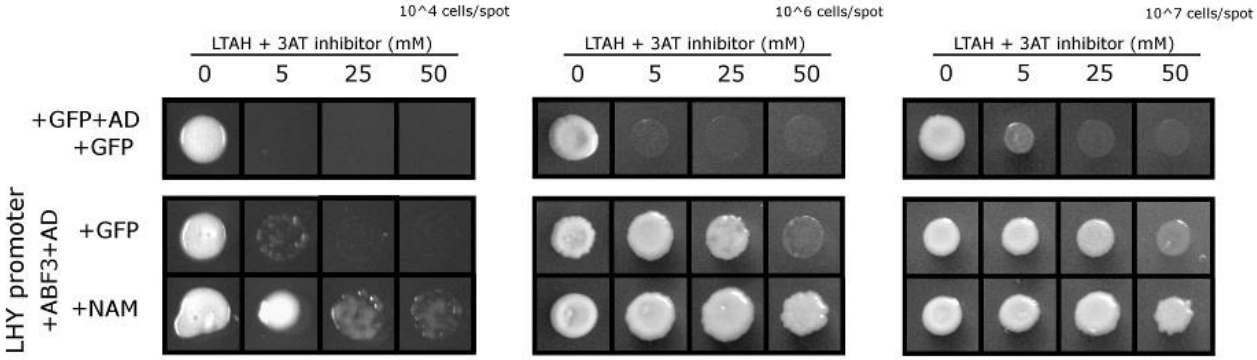
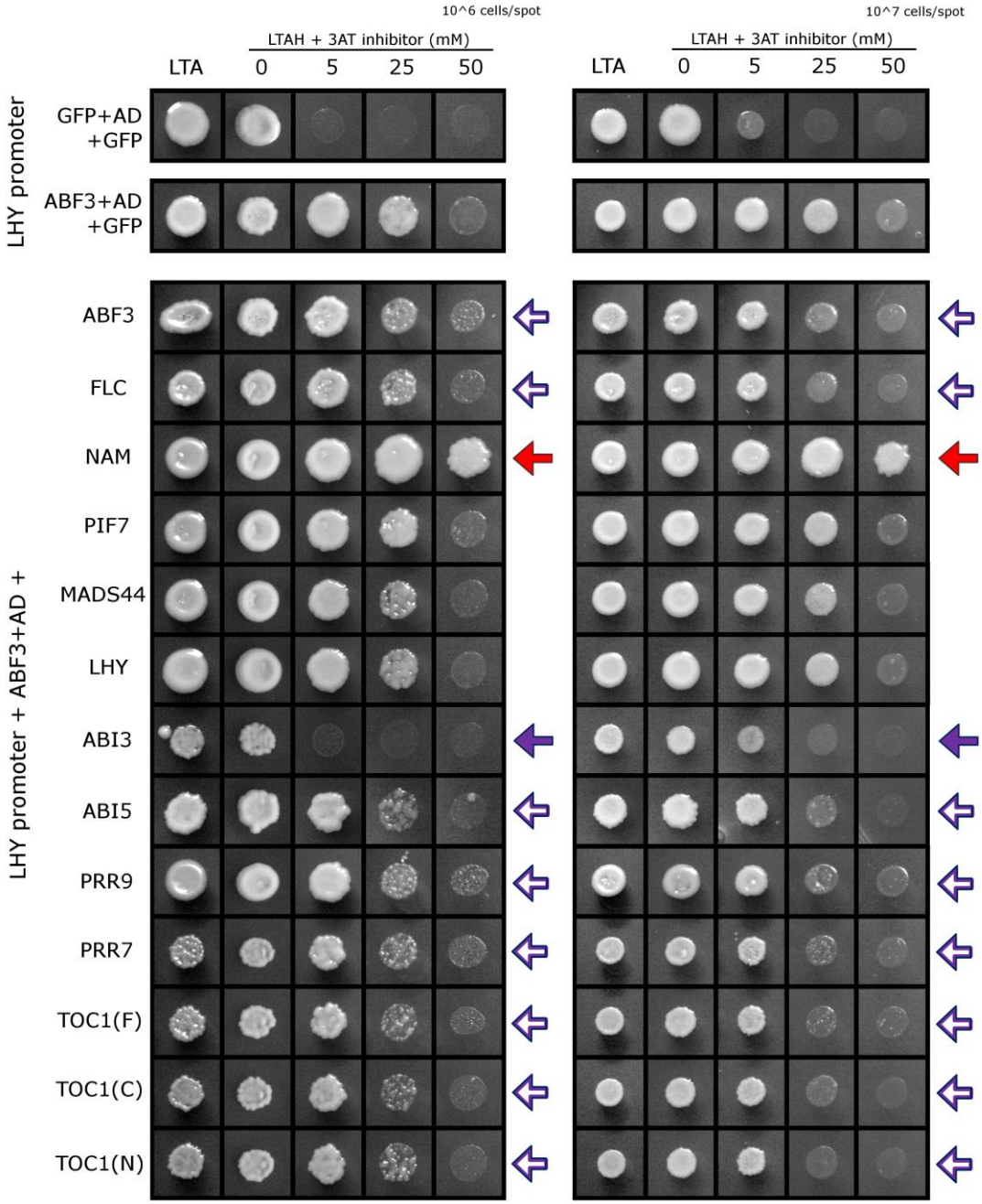
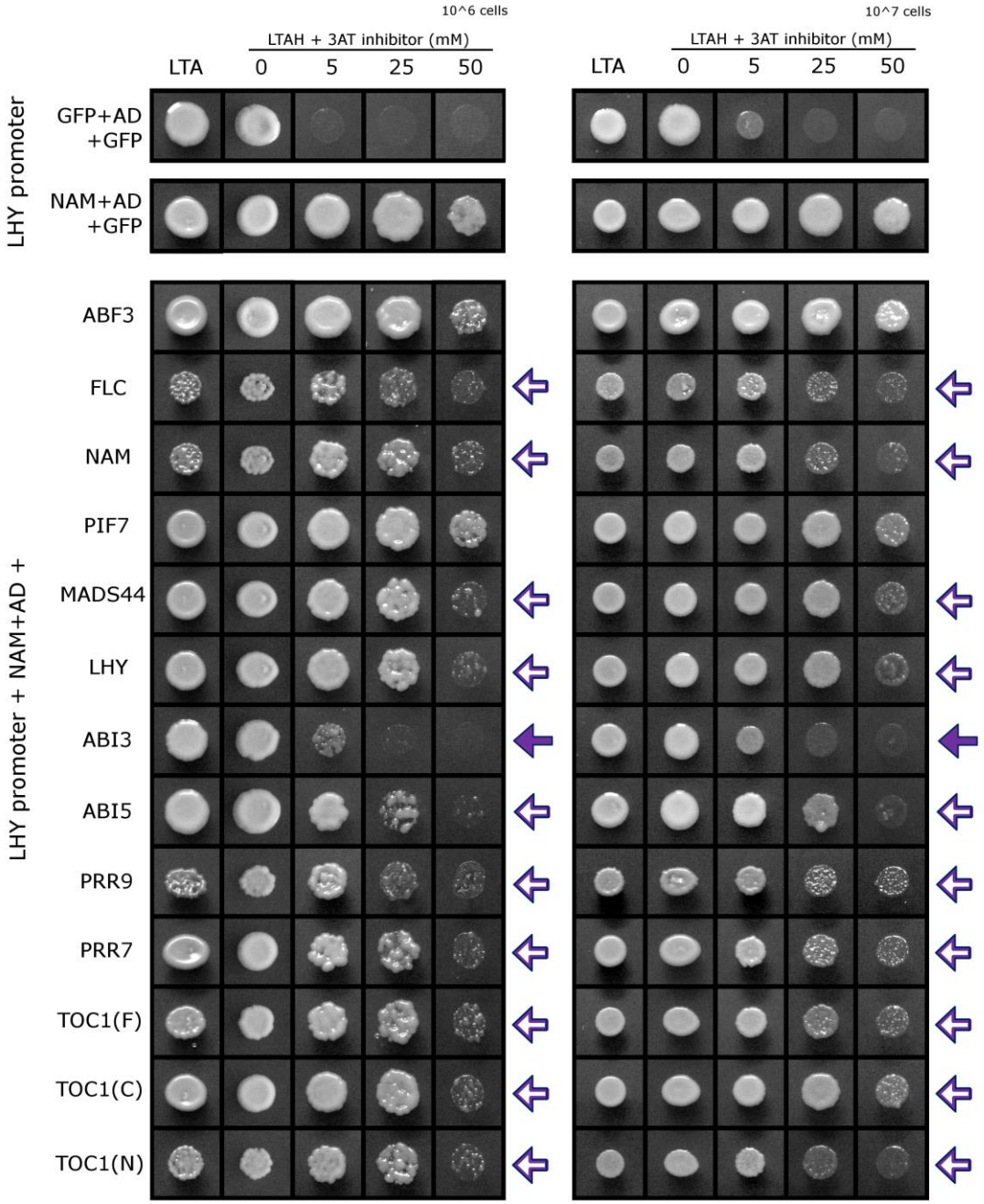


Figure 6.3: Antagonistic and synergistic interactions at the *LHY* promoter. Modified Yeast One-Hybrid assays were performed with the -957/-754 region of the *LHY* promoter in a HISLEU vector, sequentially transformed into diploid yeast with transcription factors in pDEST22 vectors (TF+AD) and pARC352 vectors (TF-AD). The activation domain (AD) is required for activation of reporter gene on binding to the promoter. Cultures were split and plated at two concentrations: 10^6 cells (left) and 10^7 cells (right) per strain per plate (B-H). Also shown in (A) are the results of a biological replicate at 10^4 cells per plate (far left). Yeast were grown for 3 days at 30°C on SD–LTA media as a control for toxicity of transcription factors to yeast, and SD-LTAH media with 0, 5, 25 or 50mM 3AT inhibitors to assay for strength of promoter binding. GFP was used as a negative control. GFP+AD + GFP is a control for auto-activation of the promoter, and any growth at this level of inhibitors is not significant. TF+AD + GFP is a test of unaided promoter binding ability of the transcription factor, therefore any growth above this level in the lower panel indicates that the second transcription factor is assisting binding to the promoter (red arrows). Growth below this level indicates inhibition of binding by the second transcription factor (purple arrows). Unfilled arrows indicate weaker results. White boxes indicate data is unavailable. A: NAM assists binding of ABF3 in biological replicate experiments, B: ABF3+AD + TFs, C: NAM+AD + TFs, D: PRR9+AD + TFs, E: PRR7+AD + TFs, F: MADS44+AD + TFs, G: FLC+AD + TFs, H: ABI5+AD + TFs.

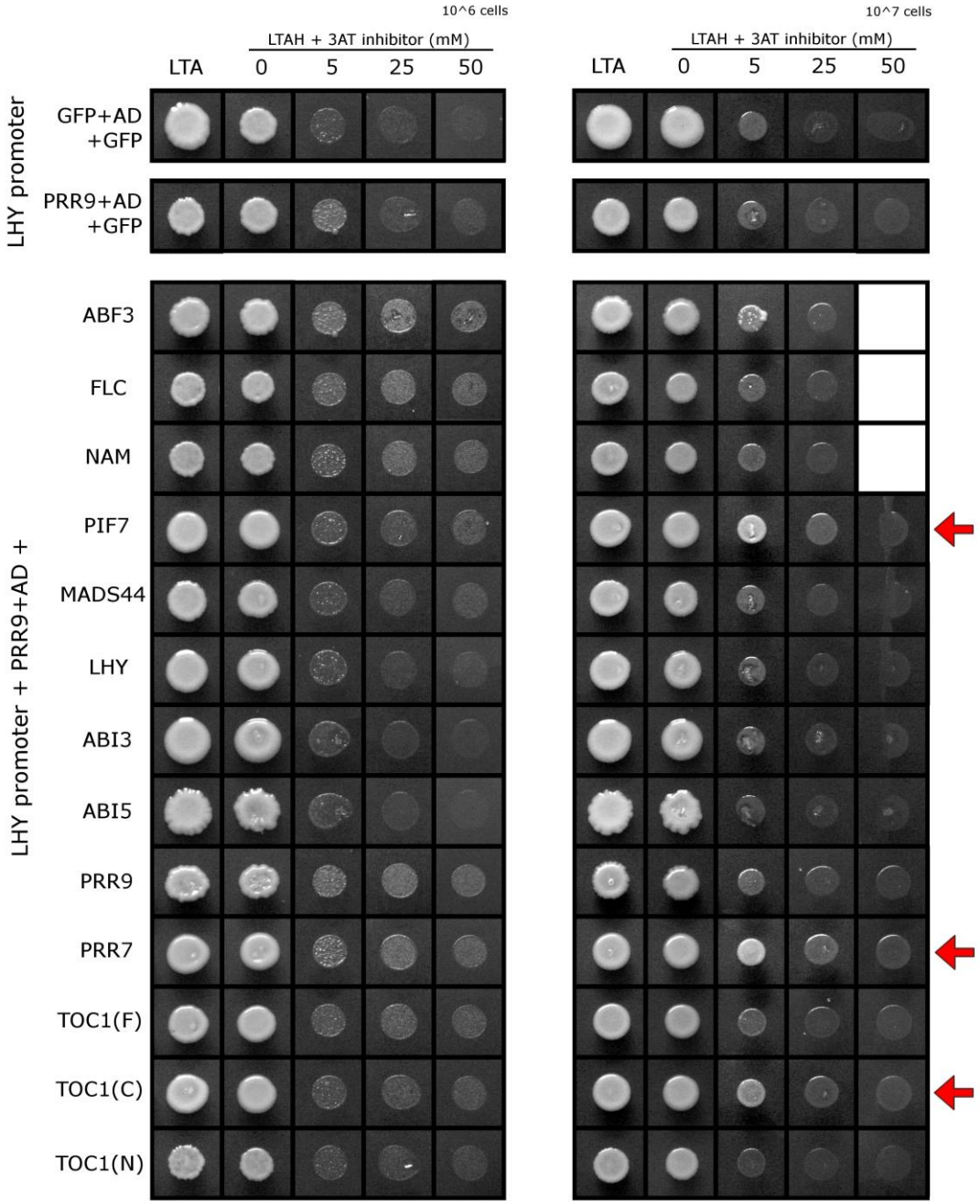
B: Binding of ABF3 to the *LHY* promoter



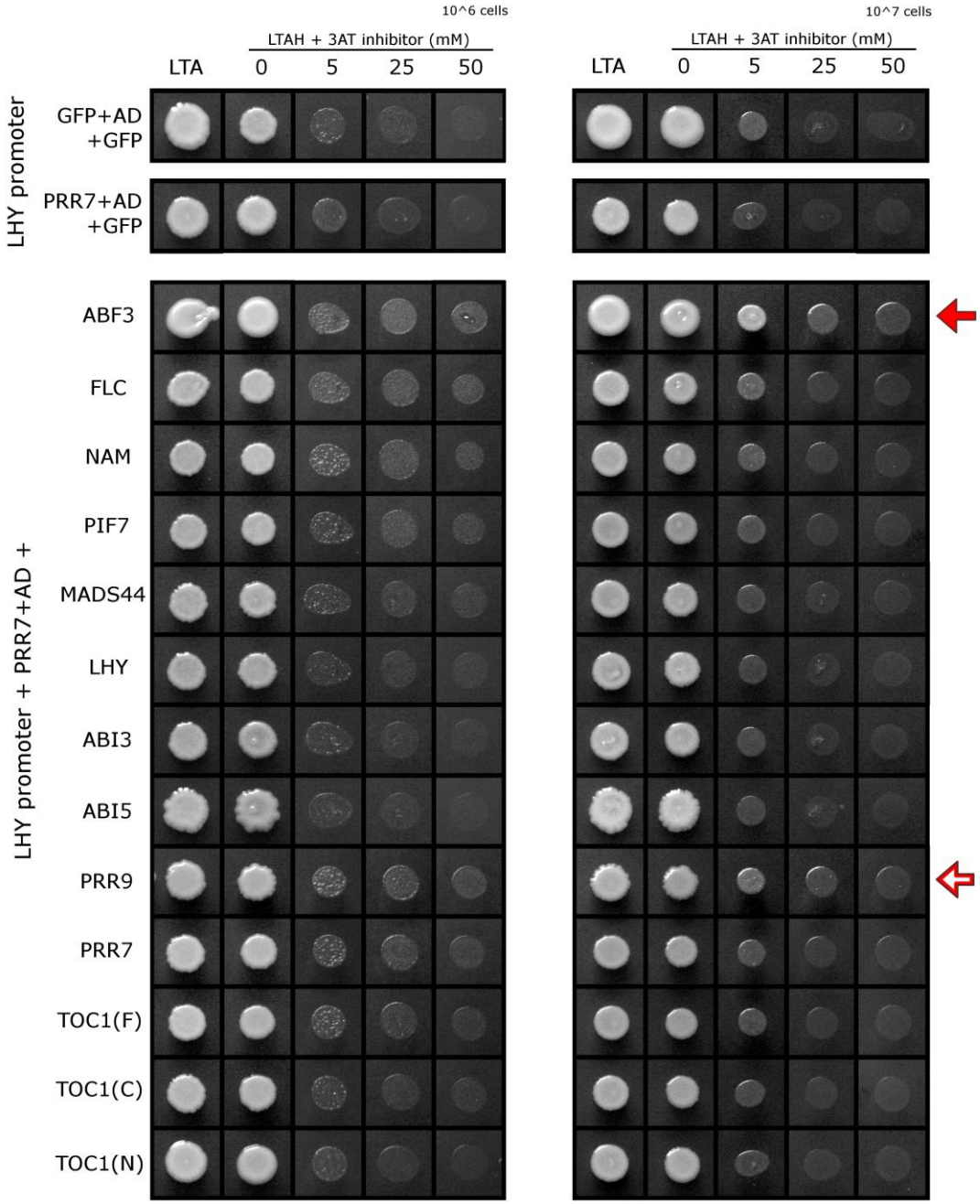
C: Binding of NAM to the *LHY* promoter



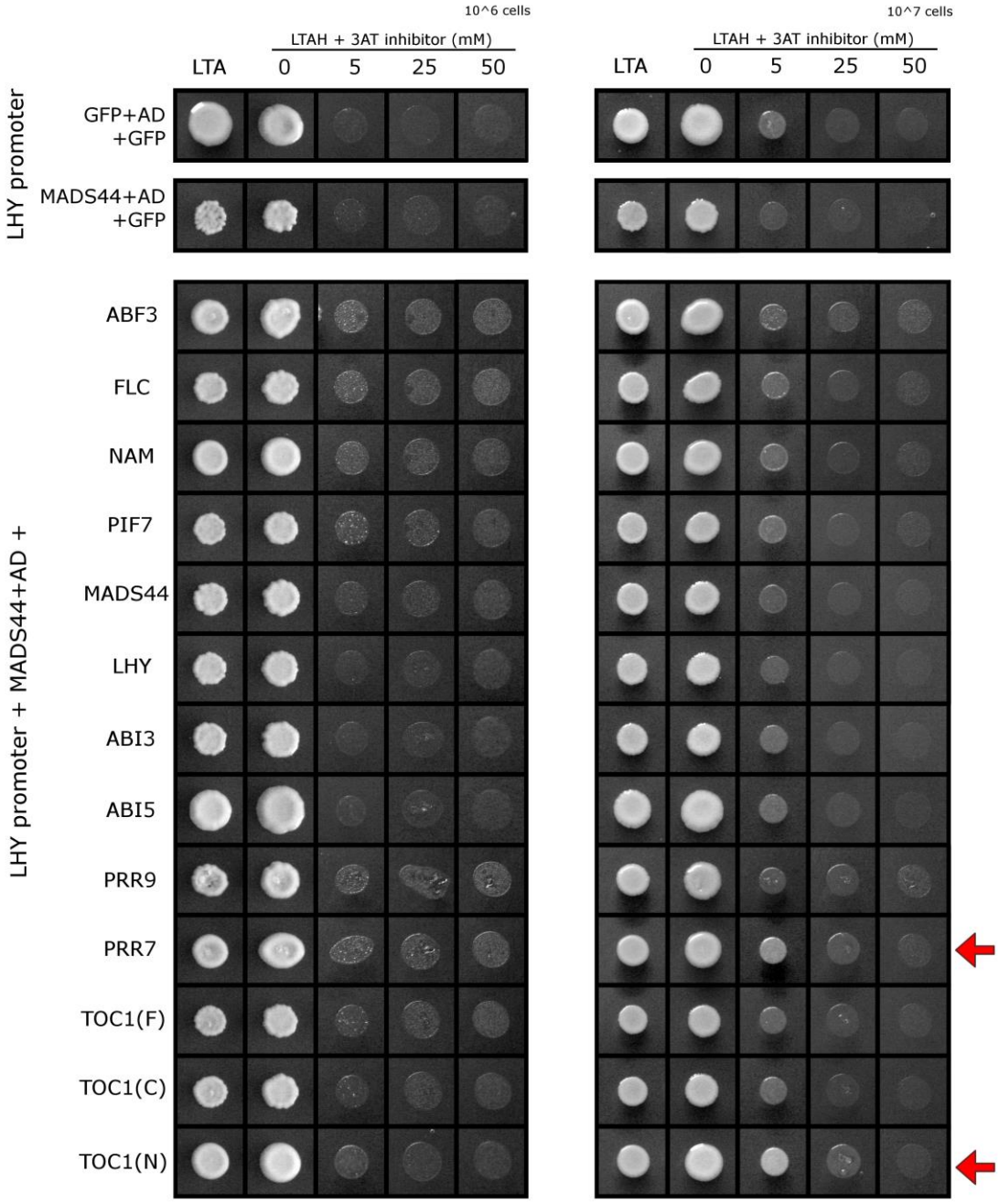
D: Binding of PRR9 to the *LHY* promoter



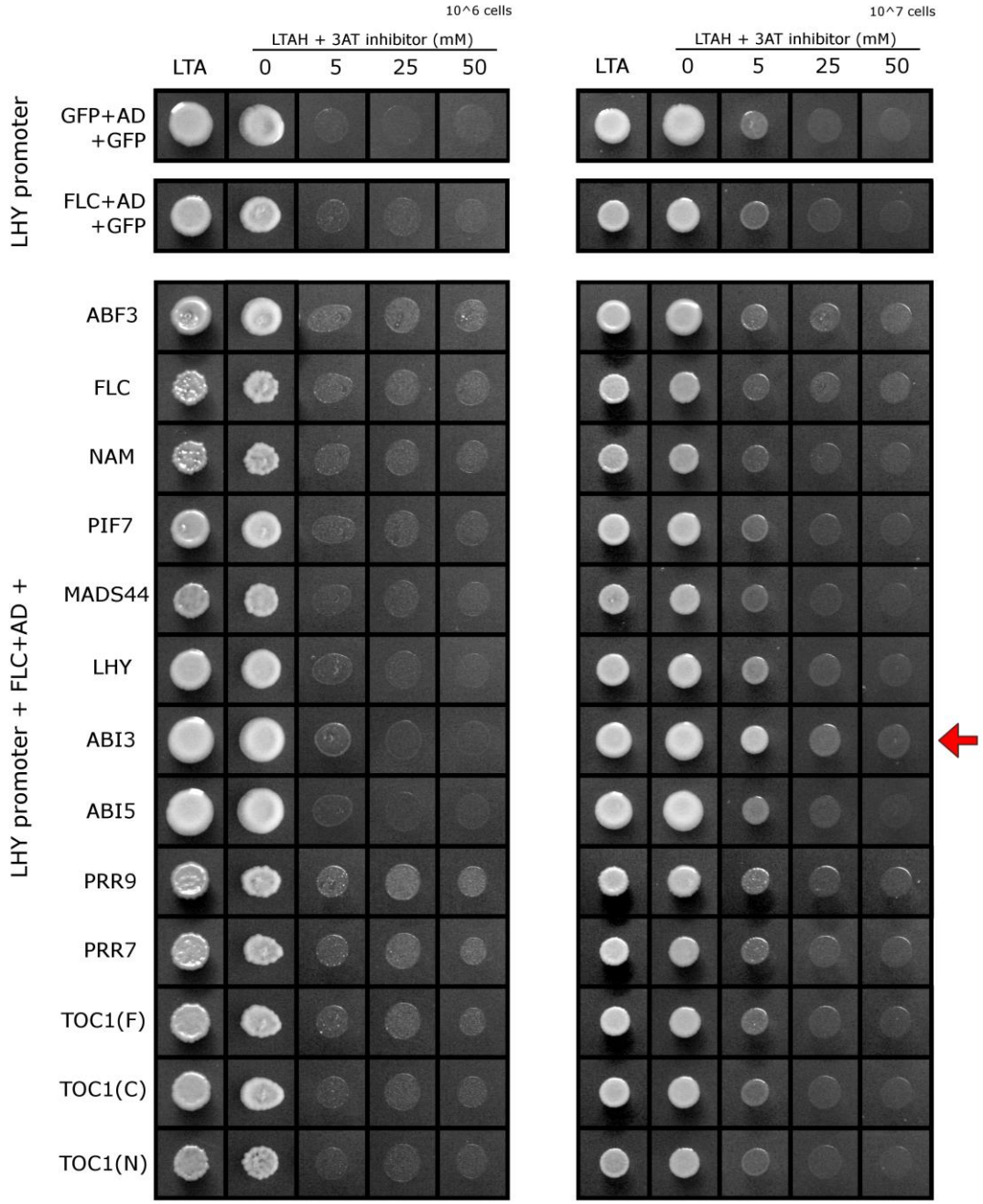
E: Binding of PRR7 to the *LHY* promoter



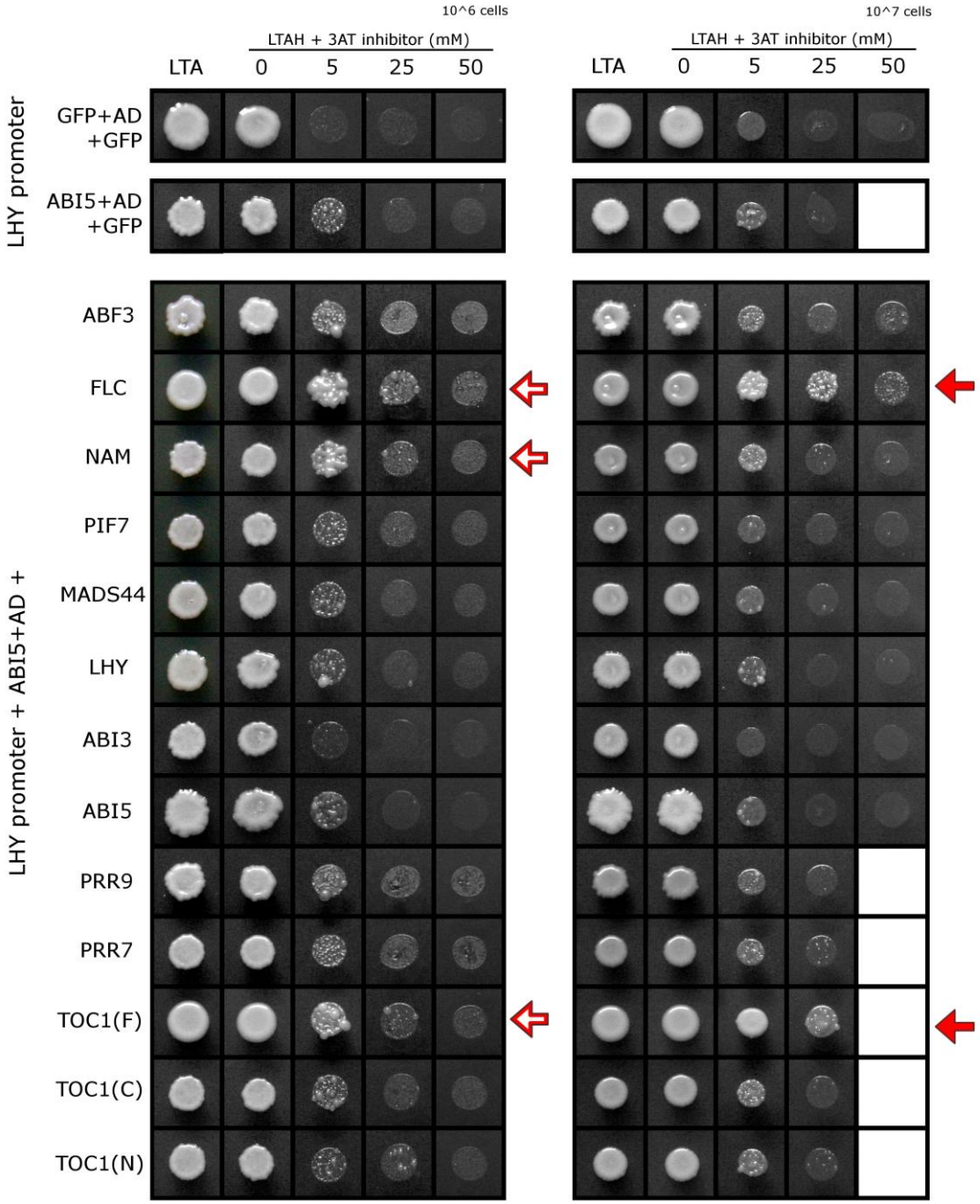
F: Binding of MADS44 to the *LHY* promoter



G: Binding of FLC to the *LHY* promoter



H: Binding of ABI5 to the *LHY* promoter



In each figure (Figure 6.3 A-H), the column labelled 'LTA' shows the growth of each yeast strain without any selection for promoter interaction. It was included as a reference and control for how well each strain grew, in case of toxicity of plant transcription factors to yeast. Despite the consistency of earlier growth rate tests, it is clear from the LTA column of Figure 6.3B that some strains grew better than others in this modified-Y1H assay. This could simply be caused by individual strains growing in non-optimal conditions during the high-density culturing of strains in 96-well plates. Alternatively, the addition of a second transcription factor may adversely affect the growth of yeast. Regardless, growth of each strain can be compared to its LTA control to assess whether apparent interaction effects are independent of differences in growth rates.

The first row of each figure (Figure 6.3 A-H) is the negative control for yeast growth on selective agar plates. It shows the growth of a strain containing the *LHY* promoter and two GFP proteins (one with and one without the activation domain). As GFP cannot bind the *LHY* promoter, any growth on LTAH selective plates (as seen in column '0') must be due to auto-activation of the promoter, as explained in Section 5.2.1.1, and any equivalent growth at this selection level in other strains should be ignored. To mask the effects of this auto-activation, yeast strains were grown on selective plates containing inhibitors of histidine biosynthesis at a range of concentrations (5mM, 25mM and 50mM). For 10^6 plated cells, any growth at 5mM inhibitors or beyond was considered significant. For 10^7 plated cells, any growth at 25mM inhibitors or increased growth at 5mM was considered significant.

The second row of each figure (Figure 6.3 A-H) shows whether the AD-tagged transcription factor is capable of binding alone to the *LHY* promoter. GFP (without AD) is also present as a negative control for the presence of an additional transcription factor-containing vector. For example, ABF3+AD + GFP in Figure 6.3B displays increased growth over the GFP+AD + GFP control and therefore ABF3 can bind the *LHY* promoter. The lower block of rows (from row 3 to 15) displays the effect of each transcription factor in turn on the binding of the AD-tagged transcription factor. For example, row six in Figure 6.3B shows the effect of PIF7 on the binding of ABF3 to the *LHY* promoter, i.e. no effect. Finally, interactions are indicated by coloured arrows: red for a TF aiding binding of TF+AD, purple for a TF inhibiting binding of TF+AD, and white with red (aiding) or purple (inhibiting) outlines for weak interactions.

6.2.2 - Results of mY1H Assays

Only two of the tested transcription factors were able to bind the -957/-754 *LHY* promoter on their own in this assay: ABF3 and NAM, and many of the other transcription factors inhibited this binding. Conversely, PRR9, PRR7, ABI5, FLC, and MADS44 were all unable to bind alone in this assay, but binding was enabled in the presence of other transcription factors. Some of the tested transcription factors did not themselves exhibit binding to the promoter under any conditions, these included full-length PIF7, LHY, ABI3 and TOC1. However, as described below, all of these were able to affect the binding of other transcription factors.

6.2.2.1 - Potential Auto-Regulation of LHY

LHY was known to bind its own promoter *in planta* from ChIP-seq data (Adams, Veflingstad and Carré, unpublished). However, LHY was unable to bind its own promoter in the mY1H assay. To investigate this discrepancy further, primers spaced along the -957/-754 region of the *LHY* promoter were used to analyse the binding pattern of LHY to its own promoter using ChIP performed on ZT-2-harvested Col-0 plants (Methods, Table 2.2). LHY was found to associate with its own promoter *in planta* in the region of the G-box (Figure 6.4). This was corroborated by ChIP on mutated promoter LHY:LUC transgenic lines, which showed reduced LHY binding when the distal 5A motifs or G-box were mutated compared to un-mutated -957/+1 *LHY* promoter (Adams, Veflingstad and Carré, unpublished).

These results suggest that LHY does in fact bind its own promoter *in planta*, and that this binding occurs at the distal 5A motifs or G-box. Therefore, the lack of LHY binding to its own promoter in yeast assays suggests that this promoter interaction requires other transcription factors not tested here. An *in vitro* assay using purified LHY protein to pull down genomic DNA provided further evidence for this suggestion, with the isolated LHY protein able to bind clock promoters including PRR7, PRR9 and TOC1, but unable to bind the *LHY* or *CCA1* promoters (Adams, Veflingstad and Carré, unpublished).

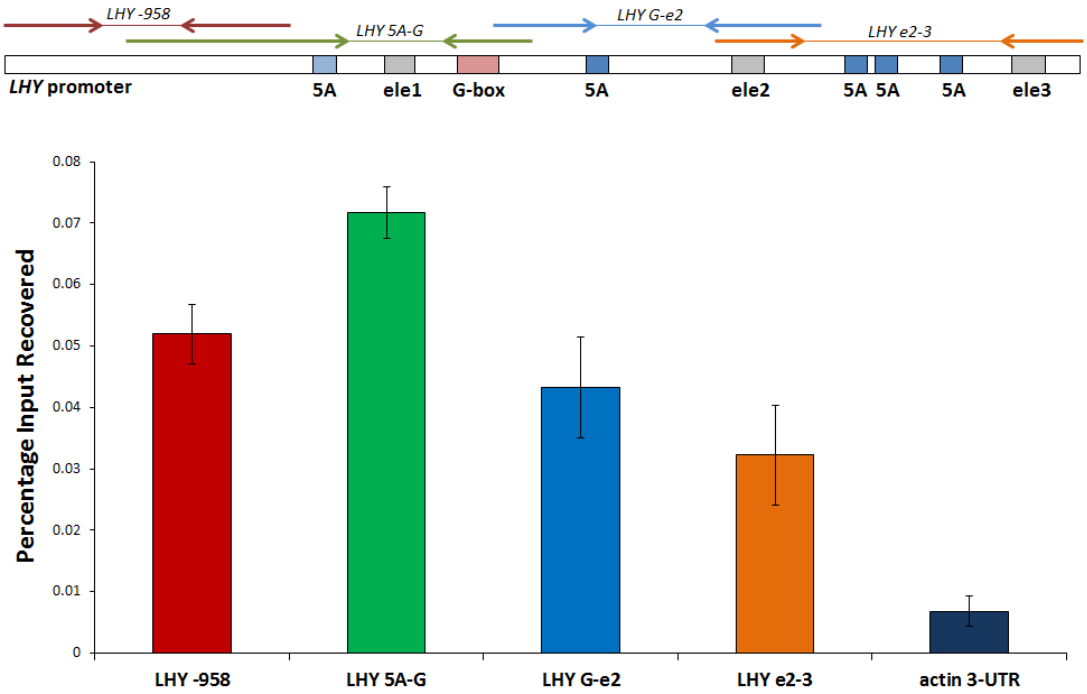


Figure 6.4: *LHY* associates with the -957/-754 region of the *LHY* promoter. In order to assay binding of LHY at different locations along the *LHY* promoter, Col-0 were grown under 12L:12D conditions at 22°C for 14 days then transferred to constant light. Tissue was harvested at ZT-2. Chromatin Immunoprecipitation (ChIP) experiments were carried out using an LHY antibody. Enrichment for *LHY* promoter sequences was tested by Q-PCR. Amplification of actin 3'UTR sequences was used as a negative control for ChIP enrichment (dark blue). Primers were designed to amplify overlapping adjacent regions in the -957/-754 region of the *LHY* promoter, as shown by the diagram above the graph. Enrichment of sequences was calculated as the percentage of Input recovered. Data shown are means of technical replicates. Error bars indicate standard deviations.

6.2.2.2 - ABA and Stress Response Factors

ABF3 was able to bind to the *LHY* promoter at 3AT inhibitor concentrations of up to 25mM (Figure 6.3B). Three transcription factors had no effect on ABF3 binding. These were PIF7, MADS44 and LHY. However, the majority of transcription factors tested inhibited ABF3 binding the promoter to varying degrees, including ABF3, FLC, ABI5, PRR9, PRR7 and TOC1 (purple arrows). This inhibition could either be by direct protein-protein interactions blocking DNA binding activity, crowding of binding sites on the promoter or competitive binding at the same site on the promoter. The inhibition of binding of ABF3+AD by non-AD-tagged ABF3

illustrates the level of yeast growth seen with competitive binding inhibition. The most striking inhibition of ABF3 binding was caused by the presence of ABI3. In this strain, growth was reduced to auto-activation levels (GFP+AD + GFP), indicating that ABI3 prevents binding of ABF3 to the promoter. Conversely, addition of NAM to the ABF3+AD strain appeared to enhance the strength of ABF3 binding, with growth clearly visible at 50mM 3AT (red arrow). This interaction was also seen in an earlier independent mY1H assay at a lower cell concentration (Exp 1, Figure 6.3A). Therefore NAM protein assists the binding of ABF3 to the *LHY* promoter.

The pathogen-responsive transcription factor NAM strongly bound to the -957/-754 *LHY* promoter fragment in this mY1H assay, with growth visible at the highest level of inhibitors (Figure 6.3C). The majority of the transcription factors tested inhibited this binding, with varying levels of effectiveness. Of these, MADS44, LHY, ABI5, PRR7, and TOC1 had the least effect on binding, all showing a similar or milder level of inhibition than the competitive binding inhibition of the NAM+AD + NAM strain. In addition, PRR9 and FLC showed inhibition of similar or greater strength than NAM+AD +NAM. Notably, PIF7 and ABF3 were the only transcription factors to not have any effect on binding of NAM to the promoter. This suggests that the mechanism by which NAM aids the binding of ABF3 (as described above) involves NAM binding to the *LHY* promoter independently of ABF3, where it can either stabilise the binding or assist the recruitment of ABF3 to the *LHY* promoter. Interestingly, ABI3 strongly inhibited NAM binding to the promoter, with growth almost back to auto-activation levels at both cell concentrations. This is analogous to

the effect of ABI3 on ABF3, suggesting that ABI3 acts to prevent binding of ABF3 both directly and by inhibiting the binding of ABF3's binding co-factors.

ABI3 was not found to bind the *LHY* promoter alone or in combination with any other protein tested in the modified-Y1H assay, and was therefore not included in Figure 6.3. ABI5 did not bind the promoter on its own (Figure 6.3H), but binding was able to occur in the presence of FLC, NAM or full-length TOC1 protein. It is particularly notable that ABI5 was able to bind to the promoter in the presence of TOC1 protein, since TOC1 can interact with the ABI5-interacting protein ABI3 in yeast (Kurup et al., 2000). Although TOC1 was unable to facilitate ABI3 binding in this assay, the link between these proteins could suggest cooperative action. This is the first indication that TOC1 may be involved in an interaction with both ABI3 and ABI5 proteins at the *LHY* promoter, potentially aiding the binding of these ABA signalling proteins.

6.2.2.3 - Flowering-Related Transcription Factors

MADS44 was also unable to bind the promoter alone in this assay (Figure 6.3F), however its binding was assisted by the presence of PRR7 and the N-terminal half of TOC1. MADS44 itself did not assist the binding of any other transcription factor tested, and had no effect on ABF3 binding. However, as described above, it did appear to inhibit the binding of NAM, suggesting either a direct protein interaction with NAM or some competitive binding to the 5A promoter motifs that are bound by NAM (Section 5.2.2.2). Again, this will require further investigation to confirm the interaction. The N-terminal half of TOC1, which is unable to bind DNA, contains the PR-domain known to enable interaction of TOC1 with other PRR proteins.

MADS44's binding being assisted by both PRR7 and TOC1-N suggests that it may be capable of interacting with the PRRs, potentially through this PR domain. However, such direct protein-protein interaction would require confirmation through Yeast Two-Hybrid assays.

Figure 6.3G displays the results of the FLC+AD mY1H assay. FLC was not able to bind the promoter in this assay, since the growth of *LHY* promoter + FLC+AD + GFP did not endure above auto-activation levels (*LHY* promoter + GFP+AD + GFP) at either cell concentration. FLC only intermittently bound the promoter in previous yeast assays. Therefore, FLC not binding here is likely an indication of specific difficulties in performing yeast assays with FLC, rather than a general inability of FLC to bind the *LHY* promoter. In support of this, the addition of ABI3 to the FLC+AD strain enables binding of FLC to the promoter. LHY may also be assisting binding of FLC+AD to the promoter. However, due to the subtlety of the effect further testing is required to confirm this interaction.

Interestingly, while ABI3 aids binding of FLC to the promoter, but is unaffected by FLC itself, binding to the promoter of the ABI3-interacting protein ABI5 is enabled by FLC, suggesting that multiple protein interactions allow for recruitment of these transcription factors to the LHY promoter. This could be speculated to occur through the known protein-protein interaction of ABI3 and ABI5, hence when ABI3 interacts with FLC to enable it to bind the promoter, FLC is brought into proximity with ABI5 which will also allow ABI5 to bind the promoter. However, this mechanism would require further study to confirm, in particular to test protein-protein interactions of FLC with ABI3 and ABI5 through Yeast Two-Hybrid assays. Regardless, it is clear

from these results that FLC, ABI3 and ABI5 are at some level involved in an interaction node at the *LHY* promoter.

6.2.2.4 - Light Signalling and Known Clock Regulators of *LHY*

PRR9 and PRR7 were also not able to bind the -957/-754 *LHY* promoter alone (Figures 6.3D and E). However, at the higher cell concentration (10^7 cells/3µl) three other transcription factors appeared to enable binding of PRR9 to the promoter: PIF7, PRR7 and the C-terminal half of TOC1. Similarly, binding of PRR7 was assisted by PRR9 and ABF3. These two temporally and functionally related transcription factors (PRR7 and PRR9) therefore appear to assist each other's binding to the *LHY* promoter. Binding of PRR9 may also be assisted by ABF3, however this result is uncertain and the interactions between ABF3 and PRR7 and PRR9 require further investigation.

Full-length PIF7 was not found to bind the *LHY* promoter alone or in combination with any other protein tested in the modified-Y1H assay, and was therefore not included in Figure 6.3. It is notable that although the full-length PIF7 was unable to itself bind the promoter, it appeared able to assist the binding of PRR9 to the promoter (Figure 6.3D). This interaction is not entirely unexpected, as PRR9 is involved in mediating light input to the clock (Farré et al., 2005), and PIF7 is part of the red light signalling pathway (Leivar et al., 2008). In addition, PIF7 is known to be capable of binding TOC1 protein. Therefore, that the C-terminal half of TOC1 also enabled binding of PRR9 to the promoter suggests a possible interaction between these transcription factors, with PIF7 and TOC1 able to interact with one another and both able to assist PRR9 binding to the *LHY* promoter. However, this

suggestion is tentative at best and requires further evidence, since this C-terminal half of TOC1, although it contains what is now known to be TOC1's DNA binding domain, does not contain the N-terminal PR domain thought to mediate TOC1's interactions with other PRR proteins (Gendron et al., 2012). This could suggest that TOC1 binding to the promoter causes a conformational change of the promoter that allows PRR9 to bind.

6.2.3 - Perturbations of the Transcription Factor Network

6.2.3.1 - *LHY* Expression in ABA Signalling Mutants

Many of the transcription factors found to bind the *LHY* promoter are known to be involved in ABA signalling. Therefore, *LHY* transcript levels were tested by Sally Adams in mutant lines for ABF3 and two other ABA-related proteins: ABI3 and ABI5 (Figure 6.5). Interestingly, ABI5 and ABI3 appear to regulate *LHY* expression at different times of day. *LHY* transcript levels in ABI5 over-expressing plants are reduced at subjective dawn and not significantly affected at subjective dusk. This suggests that ABI5 acts to repress *LHY* expression, and that this repression does not normally occur at subjective dawn. The lack of effect of overexpressing ABI5 at subjective dusk suggests that the effect of ABI5 on *LHY* expression at this time is saturated. This fits with publicly available ABI5 expression data, which shows its expression peaking around dusk in short days and under temperature cycles (Mockler et al., 2007). In addition, in *abi5-4* plants, *LHY* transcript levels were significantly increased at subjective dawn, and possibly also increased at subjective dusk. These results therefore suggest that ABI5 acts to negatively regulate *LHY* expression during the night.

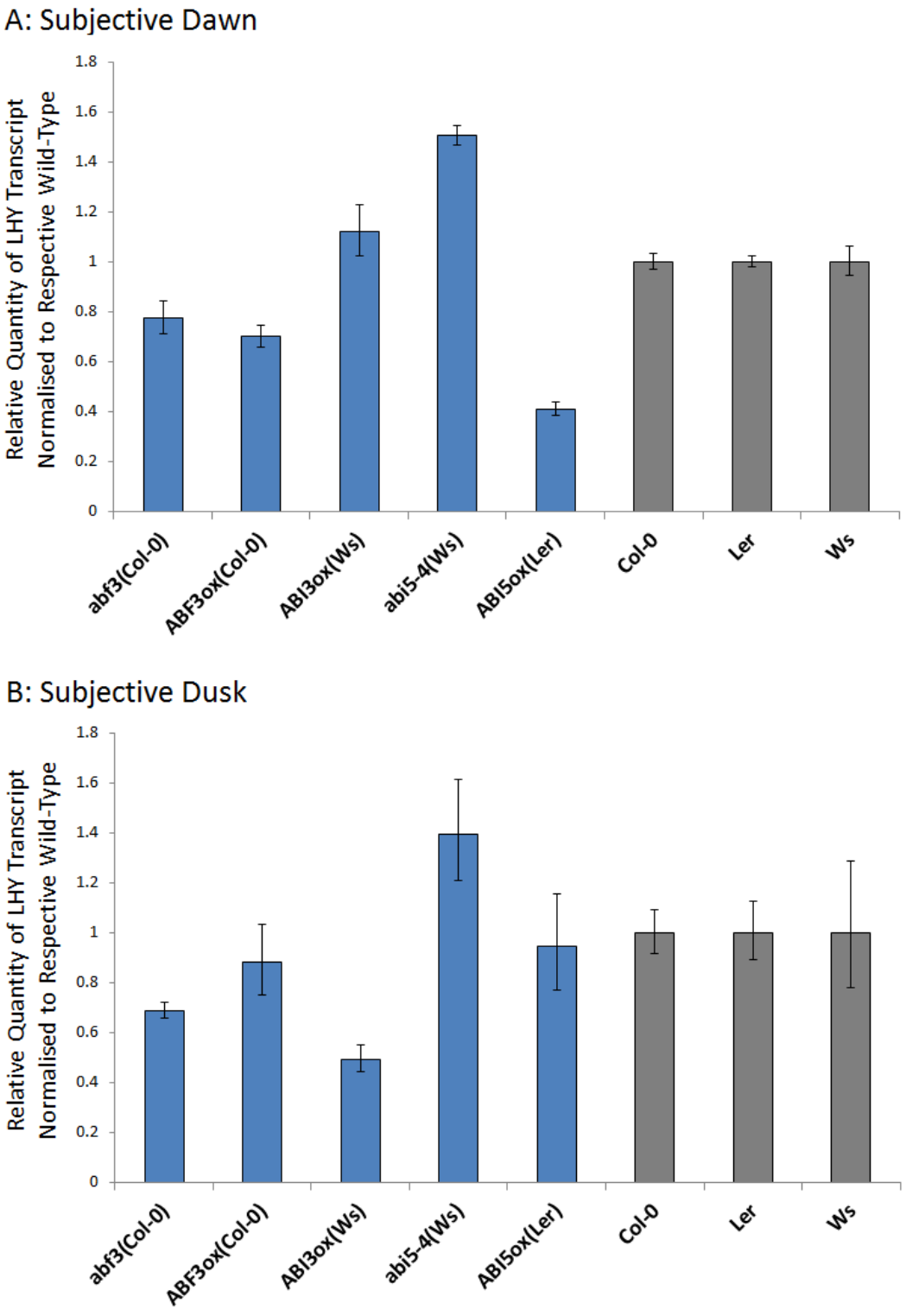


Figure 6.5: ABA signalling factors regulate *LHY* expression levels. *LHY* transcript levels in knockout and overexpressor plants (blue) relative to their respective control ecotypes (grey) at subjective dawn (ZT-24) and subjective dusk (ZT-36). Error bars indicate standard errors.

Conversely, *LHY* transcript levels are unaffected by overexpression of ABI3 at subjective dawn, but are reduced at subjective dusk, suggesting that ABI3 is not normally active at dusk, but may act around dawn to negatively regulate *LHY* expression. However, although ABI3 is known to play an important role in seed development, activating seed maturation and repressing germination (Agarwal et al., 2011), its role in wild-type adult plants has not been well characterised. In addition, ABI3 knockout lines were not available for testing, so the precise role of ABI3 in the regulation of *LHY* requires further investigation. Nevertheless, these opposing roles of ABI3 and ABI5 at different times of day, repressing *LHY* expression at dawn and dusk respectively, suggest that they might act with other time of day-dependent transcriptional co-factors to regulate *LHY* expression.

In plants over-expressing ABF3, *LHY* transcript levels were reduced at subjective dawn and unaffected at subjective dusk. This could suggest that ABF3 is required for the repression of *LHY* expression at dusk, since the reduced levels of *LHY* transcript at dawn suggest it is not normally repressing at this time. However, this does not fit with earlier predictions based on ABF3 expression data, which showed a circadian expression profile of ABF3 peaking at dawn, and a diurnal profile peaking at ZT-4. In *abf3* knockout plants, *LHY* levels were reduced at both time-points, clearly indicating that ABF3 promotes *LHY* expression. Therefore, we propose that ABF3 is an activator of *LHY* expression at dawn. However, the apparent contradiction between overexpressor and knockout results suggests that the role of ABF3 is more complex than as a simple activator.

6.2.3.2 - Effect of ABA Signalling on TOC1 Binding

As described in the section above, TOC1 was found through the modified-Y1H assay to enable binding of ABI5 to the -957/-754 region of the *LHY* promoter, and to inhibit the binding of ABF3 to this promoter region. TOC1 is also known to directly interact with ABI3 protein in yeast (Kurup et al., 2000). These interactions suggested that TOC1 has a role in regulating the action of ABA signalling factors at the *LHY* promoter. Alternatively, ABA-responsive transcription factors may modulate binding of TOC1 to the *LHY* promoter. It is therefore reasonable to speculate that TOC1's presence at the promoter may be regulated or modified by ABA signalling. Therefore, in order to determine whether the normal presence of TOC1 at the *LHY* promoter (as established by the TMG ChIP in Chapter 4) could be altered by exposure to ABA, and whether this might translate into a measurable effect on *LHY* expression, an anti-GFP TOC1 ChIP was performed on TMG plants treated with ABA. Simultaneously, the effect of ABA exposure on *LHY* expression was investigated in these same plants by extracting *LHY* RNA before and for several hours after ABA treatment.

TMG (TOC1 MiniGene) plants (Más et al., 2003a) were grown for 13 days in 12L:12D conditions, then moved to constant light on day 14. Plants were sprayed with a 25μM solution of ABA or a control solution lacking ABA at ZT-15 in the first subjective night. Samples for TOC1 ChIP were collected at ZT-16, subjected to an anti-GFP ChIP and analysed by qPCR using primers spaced along the -957/-754 region of the *LHY* promoter. Samples for *LHY* transcript level analysis were collected at ZT-15 (before spraying), ZT-16, ZT-17 and ZT-19. RNA was extracted

from these samples, and cDNA transcript levels analysed by qPCR using actin and *LHY* cDNA primers (Methods, Table 2.1).

Treatment with ABA at ZT-15 revealed a positive effect on *LHY* expression levels. Although there was only a slight increase in *LHY* transcript levels at ZT-16, 1 hour after treatment, samples taken 2 and 4 hours after treatment with ABA showed a 5-fold increase in *LHY* transcript levels relative to the control treated plants (Figure 6.6). This provides further evidence that the ABA-related proteins seen to bind the promoter in yeast have functional roles *in planta*, and indicates that ABA positively regulates *LHY* expression, with exposure to ABA causing levels of *LHY* transcript to rise prematurely during the subjective night.

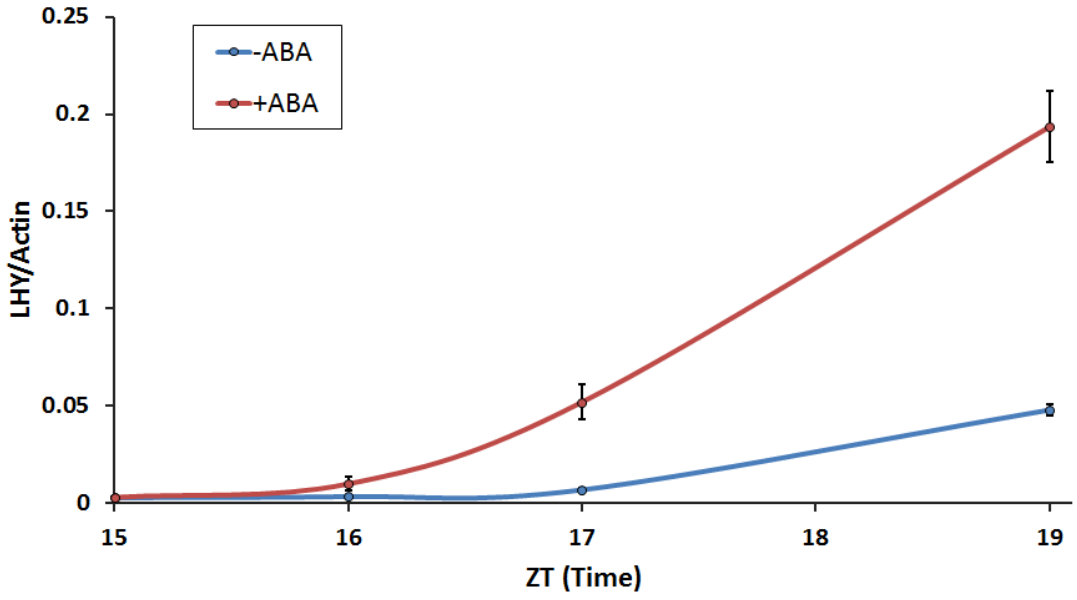


Figure 6.6: ABA treatment induces *LHY* expression. To examine the effect of ABA treatment on *LHY* expression, TMG plants (*pTOC1::TOC1:YFP* in *toc1-2* background) were grown under 12L:12D conditions at 22°C for 14 days then transferred to constant light. Plants were sprayed with a 25µM solution of ABA or a control solution at ZT-15 in the first LL cycle. Tissue was harvested before spraying (ZT-15), and at 1, 2 and 4 hours after spraying. mRNA levels were quantified from cDNA by Q-PCR using primers for LHY coding sequence (Table 2.1). *LHY* transcript levels are relative to Actin. Error bars indicate standard deviations.

However, binding of TOC1 to the promoter at ZT-16 was not significantly affected ($p>0.1$ for all primer sets) by ABA treatment at ZT-15 (Figure 6.7), suggesting that the induction of *LHY* expression by ABA may be independent of *LHY*'s regulation by TOC1.

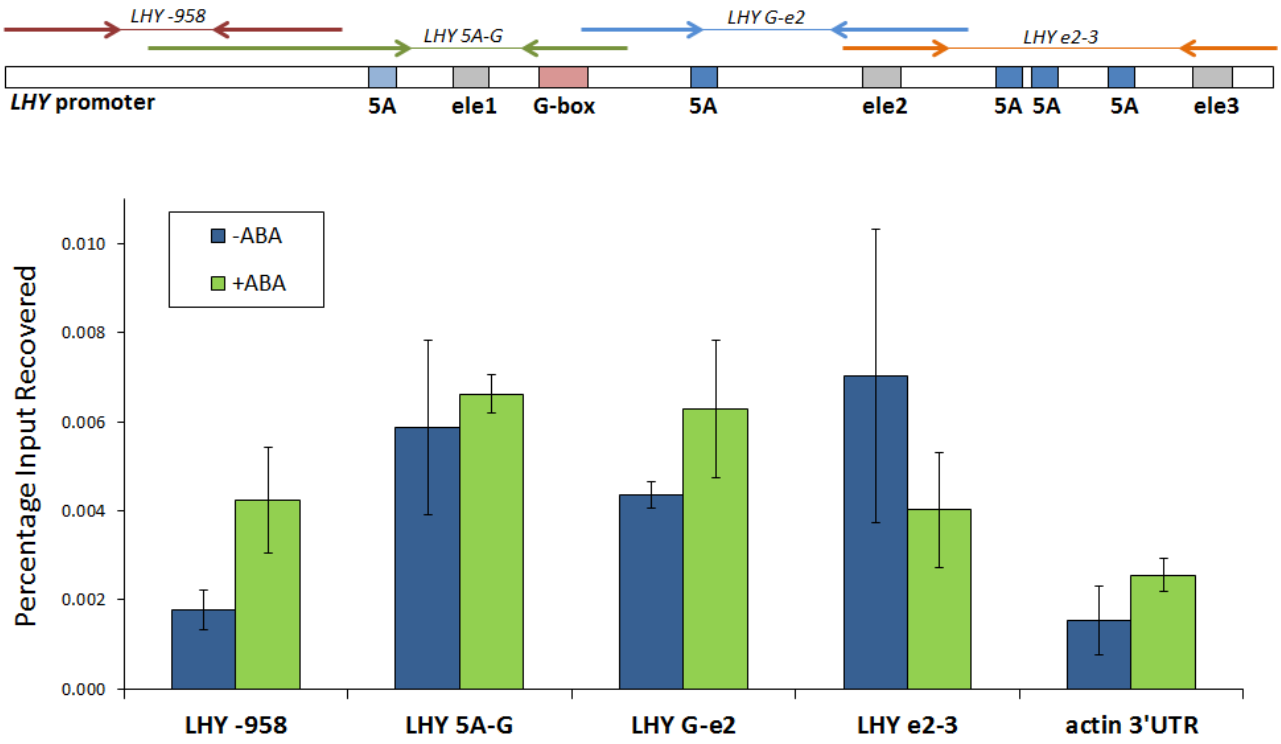


Figure 6.7: Binding of TOC1 to the *LHY* promoter is unaffected by ABA. TMG plants (*pTOC1::TOC1:YFP* in *toc1-2* background) were grown under 12L:12D conditions at 22°C for 14 days then transferred to constant light. Plants were sprayed with a 25µM solution of ABA or a control solution at ZT-15 in the first LL cycle. Tissue was harvested 1 hour later. Chromatin Immunoprecipitation (ChIP) experiments were carried out using an antibody to GFP. Enrichment for *LHY* promoter sequences was tested by Q-PCR. Amplification of actin 3'UTR sequences was used as a negative control for ChIP enrichment. Primers were designed to amplify overlapping adjacent regions in the -957/-754 region of the *LHY* promoter, as shown by the diagram above the graph. Enrichment of sequences was calculated as the percentage of Input recovered. Data shown are means of technical replicates. Error bars indicate standard deviations. Paired T-tests indicated no significant difference between ABA- and control-treated plants.

Modified-Y1H experiments indicated that ABF3's binding to the *LHY* promoter is inhibited by TOC1, ABI3 and ABI5. This suggests that ABI3 and ABI5 act alongside TOC1 as inhibitors of *LHY* expression. This role in inhibition of *LHY* expression was supported by data from ABI5 and ABI3 mutant plants (Section 6.2.3.1). However, these interactions are unlikely to occur simultaneously; although ABI5 was found to repress *LHY* at dusk, ABI3 was suggested to repress expression at dawn. It could therefore be speculated that inhibition by ABI5 occurs alongside TOC1 at night, and TOC1 aids binding of ABI3 before dawn to prevent an early accumulation of *LHY* transcript, and to allow *LHY* expression to peak sharply after dawn.

As ABF3 appears to be involved in antagonistic binding interactions with three repressors of *LHY* expression (TOC1, ABI3 and ABI5), it is most likely to have a role in activation of *LHY* expression. In addition, since TOC1 is active at night (Más et al., 2003b), and ABI5 appears to be involved in repression of *LHY* specifically at dusk, it would suggest that inhibition of ABF3 binding by these transcription factors occurs at night, perhaps to prevent premature activation of *LHY*.

These results also suggest that this antagonism may be one-way, whereby TOC1 inhibits the binding of specific ABA signalling factors, but ABA does not affect binding of TOC1 to the *LHY* promoter. Therefore, TOC1 may be a somewhat passive factor in the ABA response, assisting or preventing binding of ABA response factors only at times of day when it is already present at the *LHY* promoter. However, it should also be considered that the ChIP was performed just 1 hour after ABA treatment and that any effect of ABA on TOC1's presence at the *LHY* promoter

may not be acute enough to be detected within this timeframe. This would fit with the slow increase seen in *LHY* expression after ABA treatment, with only a minor increase in *LHY* transcript levels by ZT-16. Therefore, in order to establish whether there is a two-way antagonism between TOC1 and ABA signalling in the regulation of *LHY* expression, this ChIP experiment would have to be repeated at later timepoints when *LHY* transcript levels showed a 5-fold increase, either 2 or 4 hours after exposure to ABA.

6.3 - Discussion

6.3.1 - Summary of Results

Using a modified-Y1H assay, we were able to confirm once more that ABF3 and NAM can bind alone to the -957/-754 region of the *LHY* promoter. Notably, this assay revealed that NAM also assisted or stabilised the binding of ABF3. The binding of these two transcription factors was antagonised by the majority of other transcription factors tested, many of which are known repressors of *LHY* expression. This suggests that ABF3 and NAM may drive activation of *LHY* expression.

Earlier results had shown strong binding of the 5' truncated splice variant of PIF7 to the promoter (Table 5.1) in both pooled Y1H screens and individual assays. However, when full-length PIF7 was generated and tested in individual assays and the modified-Y1H assay, it did not bind the *LHY* promoter under any conditions. As described in Section 5.2.1.2.1, this suggests a role for alternative splicing in the light-dependent regulation of *LHY* expression. However, full-length PIF7 appeared to promote binding of PRR9 to the *LHY* promoter in the mY1H assay, suggesting a role in circadian regulation of *LHY* expression.

FLC, which had intermittently bound the promoter in previous assays, was unable to do so alone here. This is likely to be an indication of specific difficulties in using FLC in yeast assays, rather than an inability to bind the *LHY* promoter. However, FLC was able to bind in the presence of ABI3, and also interacted with ABF3 and NAM to inhibit their binding, and with ABI5 to promote its binding.

MADS44, PRR7, PRR9, TOC1, LHY, ABI3 and ABI5, which had not been identified as binding the promoter in the initial Y1H transcription factor screens, were also not found to bind alone when directly tested in this assay. However, MADS44, PRR7, PRR9 and ABI5 were able to bind the promoter in the presence of other transcription factors (MADS44: PRR7 and TOC1. PRR7: ABF3. PRR9: PRR7, PIF7 and TOC1. ABI5: FLC, NAM and TOC1). In addition, although PIF7, LHY, ABI3 and TOC1 were not able to bind the promoter either alone or with other transcription factors, they all affected the binding of other transcription factors to the *LHY* promoter. Transcription factor interactions at the *LHY* promoter are summarised in Table 6.1.

Table 6.1: Transcription factor interactions at the *LHY* promoter. Summary of interactions from the modified Yeast One-Hybrid assay, as presented in Figure 6.3.

	ABF3	NAM	MADS44	PRR9	PRR7	ABI5	ABI3	FLC	TOC1	LHY	PIF7
Binds *LHY* promoter?	yes	yes	no	No	no	no	no	no	no	no	no
Binding inhibited weakly by:	ABF3, FLC, ABI5, PRR9, PRR7, TOC1	NAM, FLC, MADS44, LHY, ABI5, PRR9, PRR7, TOC1									
Binding prevented by:	ABI3	ABI3									
Binding enhanced by:	NAM		PRR7, TOC1(N)	PRR7, PIF7, TOC1(C)	ABF3	FLC, TOC1(F), NAM		ABI3			
Binding unaffected by:	LHY, PIF7, MADS44	PIF7, ABF3									

6.3.2 - Confirmation of *in planta* Predictions in Yeast

The interaction between ABF3 and NAM provides further experimental evidence for earlier proposed regulatory interactions between different regions of the *LHY* promoter. In Chapter 3, *luciferase* reporter gene data indicated a regulatory interaction between the proximal 5A motifs and the distal promoter region *in planta*. It was proposed that the proximal 5A motifs either mediate binding of a transcriptional activator, or antagonise binding of a transcriptional repressor in the distal (G-box-containing) promoter region. Chapter 5 assigned binding sites on the *LHY* promoter to ABF3 and NAM, with NAM specifically binding the proximal 5A motifs and ABF3 binding the G-box. In the modified-Y1H assay, we have shown that NAM strengthens binding of ABF3 to the promoter. This suggests that NAM binding to the proximal 5A motifs aids activation of *LHY* expression through the stabilisation of ABF3 binding at the G-box. Therefore, through the use of yeast assays we have not only confirmed a regulatory interaction proposed from *in planta* gene expression data, but have also identified transcription factors involved.

6.3.3 - A Regulatory Transcription Factor Interaction Network at the *LHY* promoter

Through the modified-Y1H assay, we have established a number of protein interactions at the *LHY* promoter that positively or negatively affect transcription factor binding to the promoter (Table 6.1). As described below, many of the transcription factors involved in these interactions can be assigned regulatory roles, either as activators or repressors of *LHY* expression. In addition, the known abundance or activity of these transcription factors throughout the day allows us to

propose a model of changing protein interactions at the *LHY* promoter regulating *LHY* expression across the circadian cycle (Figure 6.8).

6.3.3.1 - Activation at Dawn

PIF7 is a light-stable transcription factor, involved in far-red light signalling through its interaction with PHYB (Leivar et al., 2008). It is known to target the G-boxes of light- and circadian-regulated promoters (Kidokoro et al., 2009). In the initial Y1H screen, PIF7 was found to bind the *LHY* promoter. However, this protein was a splice variant lacking the PHYB-interacting domain, and the full-length PIF7 was not found to directly bind the *LHY* promoter in this assay. However, the apparent inability of full-length PIF7 to bind the promoter cannot be taken as an indication of a biological inability to bind, as it may simply be that the full-length PIF7 requires a cofactor to bind. It is therefore expected that PIF7 would be mediating positive light signals at dawn to the *LHY* promoter.

The circadian expression of ABF3 rises through the subjective night and peaks around dawn. It is therefore likely to be an activator of *LHY* expression at dawn. This prediction was supported by *LHY* expression data in *abf3* knockout plants, which showed reduced *LHY* transcript levels at subjective dawn (Section 6.2.3). We therefore propose that ABF3 activates *LHY* expression at dawn through binding at the G-box.

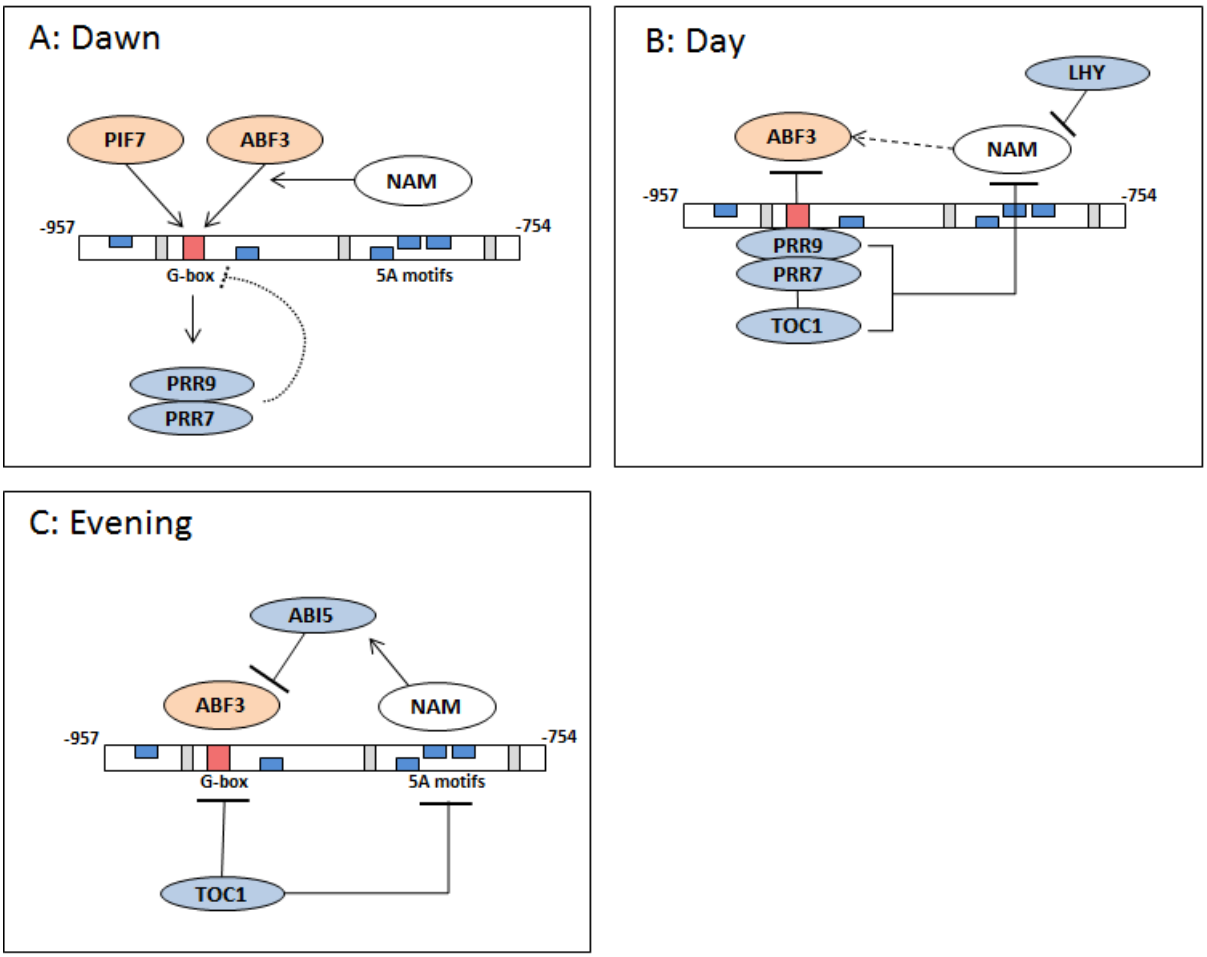


Figure 6.8: Proposed network of regulatory interactions between transcription factors acting on *LHY* expression throughout the day. Arrows between transcription factors (TFs) and promoter motifs indicate direct binding of transcription factors. Arrows between transcription factors indicate interactions assisting or stabilising binding to the promoter. Blunt ended arrows indicate inhibition of transcription factor binding. A: At dawn, PIF7 and ABF3 bind the *LHY* promoter at the G-box, and ABF3 binding is strengthened by NAM at the proximal 5A motifs, causing activation of *LHY* expression. PIF7 and ABF3 then recruit the repressors PRR9 and PRR7 to the G-box, displacing the activators. B: During the day, LHY feeds back on itself, weakening NAM binding. Waves of PRRs repress *LHY* through binding at the G-box, and inhibit ABF3 and NAM binding. C: In the evening, the PRRs are still repressing *LHY* and preventing binding of ABF3 and NAM. Any premature binding of NAM before dawn will recruit ABI5 to the promoter which inhibits ABF3 binding.

We therefore propose that PIF7 and ABF3 activate *LHY* expression at dawn through the G-box. In addition, the NAM transcription factor, which is highly responsive to pathogenic stress (Chapter 5), binds the proximal 5A motifs on the promoter and indirectly promotes the activation of *LHY* through stabilisation of ABF3's binding at the G-box. These three transcription factors therefore form an activation complex on the *LHY* promoter at dawn, and allow for fine tuning of *LHY* expression by ABA levels (via ABF3), light signals (via PIF7) and pathogen responses (via NAM). In addition, PIF7 and ABF3 can recruit PRR9 and PRR7 respectively to the promoter. These will ultimately displace the activators, as PRR7 and PRR9 inhibit binding of both ABF3 and NAM to the *LHY* promoter. This activation of *LHY* expression at dawn is therefore self-limiting (Figure 6.8A).

6.3.3.2 - Inhibition During the Day

The dawn activation module of ABF3, PIF7 and NAM on the *LHY* promoter is part of at least two negative feedback loops to constrain *LHY* expression to a narrow peak. Firstly, binding of NAM to the promoter is mildly inhibited by LHY. So the dawn activation of *LHY* expression will quickly result in weakening of NAM binding as LHY protein levels rise. Secondly, as described above, PIF7 can recruit PRR9 to the promoter and ABF3 can recruit PRR7. In addition, PRR9 and PRR7 will assist each other's binding. Once at the promoter, PRR9 and PRR7 antagonise the binding of ABF3 and NAM, likely in part through direct displacement from the promoter, since PRR9 and 7 are thought to target the G-box-containing region of the *LHY* promoter (Nakamichi et al., 2010). This displacement will be dependent on the timing of expression of PRR9 and PRR7, which have their transcriptional peaks at ZT-2 and ZT-6 respectively (Matsushika et al., 2000).

In the early evening, when it is present, TOC1 will also inhibit ABF3 and NAM binding. Although PRR5 was not tested in this assay, these results suggest that the PRRs (PRR9, 7, 5 and TOC1), which are sequentially expressed throughout the day (Matsushika et al., 2000), act to antagonise the binding of these activators at the *LHY* promoter during the day and into the evening. Furthermore, the PRRs are now known to be direct repressors of *LHY* expression (Nakamichi et al., 2010). Therefore, the dawn activators help recruit repressors to the promoter, which will both actively repress *LHY* expression and inhibit activator binding throughout the day and evening (Figure 6.8B).

This suggests a possible explanation for the apparently contradictory effects on *LHY* expression in ABF3 mutant plants described in Section 6.2.3.1. Although *abf3* knockout plants showed reduced levels of *LHY* expression, indicating activation by ABF3, a reduction in *LHY* transcript levels was also seen at subjective dawn in ABF3 overexpressing plants. This suggested that ABF3 was also required for repression of *LHY*, and that this repression was likely to occur in the evening, since it did not normally occur at dawn. The modified-Y1H data shows that ABF3 promotes recruitment of the repressor PRR7 to the *LHY* promoter, and that the PRRs then promote each other's binding to the promoter. Therefore, by recruiting PRRs to the *LHY* promoter, ABF3 is indirectly promoting repression of *LHY* later in the day.

6.3.3.3 - Inhibition at Night

TOC1 is present throughout the night (Matsushika et al., 2000), and will be both repressing *LHY* expression and inhibiting early binding of ABF3 and NAM at this time. However, an additional evening inhibitor of *LHY* expression was identified

through the modified-Y1H assay. The binding of ABI5 to the *LHY* promoter was enabled by TOC1, and both TOC1 and ABI5 inhibited the binding of ABF3 and NAM. Expression of ABI5 peaks around dusk and in the evening under temperature (HC) cycles and short days, around ZT-12 and ZT-8 respectively, with the trough of expression around dawn (Mockler et al., 2007). The timing of ABI5's expression and its interaction with TOC1 suggested a role for ABI5 in the repression of *LHY* expression (Figure 6.8C). Analysis of *LHY* transcript levels in ABI5 overexpressor and mutant plants confirmed a functional role for ABI5 in the repression of *LHY* expression at night (Section 6.2.3).

Notably, NAM also assists binding of ABI5 to the *LHY* promoter, and ABI5 inhibits NAM binding. This may be another mechanism to prevent early activation of *LHY* expression, whereby NAM will antagonise its own binding by recruiting ABI5 if it binds the promoter too early before dawn. Alternatively, since expression data suggests that NAM may be expressed at a low level throughout the day (Mockler et al., 2007), NAM could have a dual role in aiding the binding of both activators and repressors depending on the time of day.

6.3.4 - Impact of Plant Development and Environmental Stresses on the Transcription Factor Interaction Network at the *LHY* Promoter

6.3.4.1 - Impact of Plant Development: Germination and Flowering

The ABA signalling protein ABI3 is well-established as being an essential factor in late embryogenesis and the regulation of seed dormancy. Concordantly, ABI3 is known to primarily target the promoters of genes involved in seed maturation,

although it has also been found to target ABA-regulated genes and those involved in seedling development and freezing responses (Mönke et al., 2012).

ABI3 was not found to bind the *LHY* promoter in the modified-Y1H assay, and binding was not enabled by the presence of other transcription factors tested. This is unsurprising as the *LHY* promoter lacks the RY motif targeted for binding by ABI3 (Mönke et al., 2004). However, ABI3 strongly inhibited binding of both ABF3 and NAM to the promoter, to the extent of reducing yeast growth to auto-activation levels. This suggests that ABI3 prevents binding of these transcription factors to the *LHY* promoter. The inhibition of ABF3 binding to the G-box is particularly notable since the G-box related GBL (ACGTG(T/G)C) motif was recently found to be over-represented in the promoters of ABI3 target genes (Mönke et al., 2012). In addition, ABI3 has also been recently shown to interact with the G-box-targeting phytochrome interacting factor 3-like 5 (PIL5) to activate expression of a negative regulator of germination (Park et al., 2011, Oh et al., 2009).

However, it is unknown whether ABI3 has a role in adult plants, being primarily involved in embryogenesis and the regulation of seed dormancy and germination. This suggests that the action of ABI3 on the *LHY* promoter may be restricted to this early developmental time-frame, which includes the regulation of seed dormancy and initiation of the circadian clock (Penfield & Hall, 2009).

The flowering regulator FLC had been previously established as targeting the *LHY* promoter *in planta* (Spensley et al., 2009). Through the modified-Y1H assay, FLC was found to both inhibit binding of the dawn activation module of ABF3 and NAM

to the *LHY* promoter, and also to promote binding of the evening-specific *LHY* repressor ABI5. Although FLC was not found to bind to the *LHY* promoter in this yeast assay when alone, the presence of ABI3 enabled binding of FLC to the promoter. These interactions suggest that FLC is part of the inhibitory machinery at the *LHY* promoter, resulting in repression of *LHY* expression when FLC is present. This correlates with FLC's known role as a transcriptional repressor (Helliwell et al., 2006).

The related MADS-box transcription factor, MADS44, also appears to have a role in these inhibitory interactions. MADS44 was unable to bind the promoter alone, but it was able to bind in the presence of some PRRs. MADS44 in turn inhibited the binding of NAM, suggesting that it antagonises the binding of activators after being recruited to the *LHY* promoter by repressors of *LHY*. MADS44 has been previously shown to interact in yeast with the flowering regulator SOC1, which is directly repressed by FLC (de Folter et al., 2005, Lee & Lee, 2010). These results suggest that developmental signals involved in the regulation and initiation of flowering can also feed into the *LHY* promoter, potentially altering *LHY* expression.

6.3.4.2 - Impact of Environmental Stresses

The NAM transcription factor is highly responsive to pathogenic stress, with expression at a low level in most tissues until the plant is exposed to pathogens or related markers of infection (Chapter 5). Possibly as a result of this, its diurnal expression profile is variable. In most available studies, NAM is present throughout the day and night at a constant level. However, it can also display strong rhythmic expression under 12L:12D conditions (Col-0), with a transcriptional peak just after

dawn (DIURNAL, Mockler et al., 2007). The effect or degree of influence of NAM is therefore highly likely to be affected by pathogenic stress. In addition, NAM was found to assist binding of the ABA signalling factors ABF3 and ABI5 to the *LHY* promoter in the modified-Y1H assay. Binding of NAM was also prevented by another ABA-related factor, ABI3. Since the hormone ABA is well-known to mediate stress responses, this interaction of ABA signalling factors at the *LHY* promoter with the pathogen-induced transcription factor NAM suggests that the expression of *LHY* is altered in response to pathogenic stress.

Notably, there is also evidence of a positive feedback loop between the clock and the ABA signalling pathway: ABA was found to induce *LHY* expression within 2 hours of treatment at ZT-15 (Figure 6.6), and *ABF3* expression in *LHY* over-expressing plants is elevated at its transcriptional peak (Mockler et al., 2007). Such a positive feedback loop would enable rapid clock-mediated amplification of the ABA signal, potentially in response to sudden abiotic environmental stresses. In addition, the up-regulation of NAM in response to pathogen attack would augment this positive feedback loop, and suggests that this method for signal amplification could also be used in plant defence.

CHAPTER 7

Summary of Results and Future Work

7.1 - Research Aims

The research presented in this thesis aimed to further investigate the regulatory roles of known motifs within the *LHY* promoter, to identify transcription factors binding these motifs and to explore how interactions between these transcription factors and TOC1 and the PRRs might serve to regulate *LHY* expression.

7.2 - A Complex Regulatory Network of Transcription Factor Interactions at the *LHY* Promoter Enables Integration of Light and Environmental Stress Signals to the Clock

The roles of evolutionarily conserved motifs within the *LHY* promoter were investigated through *luciferase* assays (Chapter 3). The CT-rich region was found to regulate rhythmic expression of *LHY*, the first motif on the promoter to be identified as such. The CT-rich region also activated expression redundantly with the distal promoter region. Two nodes of activation were therefore identified on the promoter, with one able to compensate for loss of the other. In addition, the proximal 5A motifs were found to interact with the distal promoter region to mediate activation. A general requirement for 5A motifs within the promoter for activation of *LHY* expression was also established.

The clock protein TOC1 was also found to be required for expression of LHY:LUC reporter constructs (Chapter 4), and a link between TOC1 and the CT-rich region was suggested due to the similarity of these results. However, this interaction could

not be through direct binding, as TOC1 was found to bind further upstream, around the G-box. It was therefore proposed that TOC1 interacts with the CT-rich region indirectly, either through CT-rich region mediated recruitment of TOC1 to the promoter, or TOC1-dependent regulation of transcription factors targeting the CT-rich region. In addition to directly binding the G-box, TOC1 was proposed to be indirectly required for activation at the G-box, through repression of another G-box targeting negative regulator of *LHY*, such as PRR9.

There was also evidence for the presence of TOC1-independent light-induced activation of *LHY* in the distal region of the promoter (Chapter 4). This is likely to occur through the G-box, since it is known to play a role in the mediation of light signals (Martinez-Garcia et al., 2000). The identification of GBF1 and PIF7, both G-box targeting light signalling factors, as binding the *LHY* promoter in yeast suggests that we may have identified at least some of the transcription factors responsible for this light-induced activation.

Fifteen transcription factors were identified as binding the -957/-754 region of the *LHY* promoter from the Yeast One-Hybrid screen (Chapter 5). These included transcription factors with roles in light signalling (GBF1, PIF7), biotic and abiotic stress responses (ABF3, ABF4, EEL, DPBF2, NAM), photomorphogenesis (IAA2) and flowering (ANT, FBH2). Many of these transcription factors had potentially overlapping roles, particularly in regard to hormone signalling pathways associated with development and stress responses including Auxin, Abscisic Acid (ABA), Brassinosteroids (BR) and Salicylic Acid (SA). In addition, *in planta* data suggested that some of these transcription factors were possible activators of *LHY* transcription,

including bZIP52, ABF3 and GBF1. If confirmed, these would be the first activators to be identified as targeting the *LHY* promoter.

The high incidence of ABA signalling and responsive transcription factors in the collection found to bind the *LHY* promoter, led to an investigation of the effect of the ABA signalling pathway on *LHY* expression. ABA was able to induce *LHY* expression during the night, suggesting that it has a generally positive influence on LHY. However, measurement of *LHY* transcript levels in ABF3, ABI5 and ABI3 knockout and overexpressor plants revealed distinct regulatory roles for these proteins. ABI3 and ABI5 were suggested to repress *LHY* during the day and night respectively. ABF3, which is itself induced by ABA, was implicated in both activation and repression of *LHY*, suggesting a complex role for ABF3 that requires further investigation.

Specific binding sites on the *LHY* promoter were identified for several of these transcription factors, in both proximal and distal regions of the promoter (Chapter 5). Within the proximal promoter region, the $5A_{345}$ motifs were shown to be a specific binding target for NAM, and ANT bound specifically to Element 3. Within the distal promoter region, ABF3, GBF1 and bZIP29/30 all bound the G-box, and this binding appeared to be stabilised by G-box flanking elements including the 5A motifs and the novel Elements 1 and 2. These results provide the first evidence that these evolutionarily conserved novel motifs (Elements 1-3) have a functional role within the *LHY* promoter.

A number of transcription factor interactions at the *LHY* promoter were identified through a modified Yeast One-Hybrid assay (Chapter 6). These comprised both antagonistic interactions, such as inhibition of ABF3 binding by PRR9, PRR7 and TOC1, and synergistic interactions, such as the stabilisation of ABF3 binding by NAM, and of PRR7 by PRR9. These results suggested once more that ABF3 is an activator of *LHY* expression, since its binding is antagonised by known repressors of *LHY*. Importantly, these yeast-based results fit well with the conclusions from the *luciferase* experiments, chiefly that the proximal 5A motifs assist activation within the distal promoter region. This *in planta* interaction was mirrored in the yeast assays when NAM, bound at the proximal 5A motifs, was shown to assist binding of the probable activator ABF3 at the G-box.

In addition, these yeast results established interactions between clock and ABA signalling transcription factors in the regulation of *LHY* expression, with known repressors of LHY such as TOC1 and the PRRs acting in parallel with the ABA signalling factors ABI3 and ABI5 to inhibit binding of the probable activation module of ABF3 and NAM. Furthermore, a positive feedback loop was proposed between ABF3 and LHY, whereby the clock could be appropriated for the rapid amplification of stress signals in response to sudden abiotic or pathogenic stresses.

7.3 - Possible Future Work

The research presented in this thesis offers new insights into the complex transcription factor interactions that occur to regulate the *Arabidopsis* circadian clock gene *LHY*, and how these may facilitate integration of entrainment and stress-induced signals to the clock. The identification of transcription factors with roles in

so many different plant processes as binding to the *LHY* promoter has opened up a wide array of new avenues for further research. Some possible experiments to clarify and expand on these findings are described below.

- The novel Elements 1-3 in the *LHY* promoter are now known to have functional roles in the mediation of transcription factor binding. However, the precise role of these elements in the regulation of *LHY* expression remains unknown. This could be investigated by site-directed mutagenesis of these motifs in LHY:LUC constructs using the primers designed for mutagenesis of Elements 1-3 in the LHY:HISLEU yeast reporter construct. *Luciferase* assays with these reporter constructs would enable elucidation of the role of these elements in the regulation of *LHY*'s rhythmic expression, as illustrated for the CT-rich region in Chapter 3.

- Additional analysis of *LHY* transcript levels in mutant plants, as described in Section 5.2.1.3, is required to confirm the regulatory effects of transcription factors found to bind the *LHY* promoter. In addition, LHY:LUC constructs could be introduced into the homozygous SALK lines already obtained to examine the effect of each transcription factor on temporal expression patterns of the reporter gene.

- Although fifteen transcription factors were identified as binding the *LHY* promoter, the binding of just six of these was mapped to specific promoter motifs. The mutated promoter Yeast One-Hybrid assay could therefore be repeated with freshly transformed yeast strains to avoid any problems of sub-

optimal growth. In addition, individual assays for these transcription factors could be performed by adapting the multimers originally constructed for motif sufficiency *luciferase* tests (Chapter 3) for use in yeast assays.

- Since the majority of transcription factors identified from the Yeast One-Hybrid screen were not tested for interactions, the modified-Y1H assay could be repeated with different combinations of transcription factors. GBF1 in particular would be an interesting target for this experiment, since full-length PIF7 was not able to bind to the promoter in the assay, and so investigation of interactions between clock and light signalling proteins was limited.

- The role of ABA in the clock-mediated regulation of *LHY* transcription could be investigated further. For example, although ABA had no effect on TOC1 binding in the TMG ChIP, does it affect the binding of LHY or any other transcription factor to the *LHY* promoter?

- The effect of Auxin on the regulation of *LHY* expression could also be investigated through qPCR analysis of *LHY* transcript levels after Auxin treatment, or the introduction of LHY:LUC constructs into *iaa2* plants. Since so many of the transcription factors identified as binding the *LHY* promoter are differentially regulated in response to pathogenic stress, similar experiments could also be performed using pathogens or bacterial elicitors to examine the role of the plant defence response in the regulation of *LHY* expression.

REFERENCES

Abdulrehman D, Monteiro PT, Teixeira MC, *et al.*, 2011. YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface. *Nucleic Acids Res* 39, D136-D40.

Abel S, Oeller PW, Theologis A, 1994. Early auxin-induced genes encode short-lived nuclear proteins. *Proc Natl Acad Sci U S A* 91, 326-30.

Adams S, Carré IA, 2011. Downstream of the plant circadian clock: output pathways for the control of physiology and development. *Essays Biochem* 49, 53-69.

Agarwal P, Kapoor S, Tyagi AK, 2011. Transcription factors regulating the progression of monocot and dicot seed development. *Bioessays* 33, 189-202.

Alabadi D, Oyama T, Yanovsky MJ, Harmon FG, Más P, Kay SA, 2001. Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* 293, 880-3.

Alabadi D, Yanovsky MJ, Más P, Harmer SL, Kay SA, 2002. Critical role for CCA1 and LHY in maintaining circadian rhythmicity in Arabidopsis. *Curr Biol* 12, 757-61.

Amasino R, 2010. Seasonal and developmental timing of flowering. *The Plant Journal* 61, 1001-13.

Barnes MR, Gray IC, 2003. *Bioinformatics for Geneticists*. First Edition. Wiley Publishing, UK.

Bell-Pedersen D, Cassone VM, Earnest DJ, *et al.*, 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet* 6, 544-56.

Bensmihen S, Giraudat J, Parcy F, 2005. Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during Arabidopsis thaliana embryo maturation. *J Exp Bot* 56, 597-603.

Bensmihen S, Rippa S, Lambert G, *et al.*, 2002. The homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. *Plant Cell* 14, 1391-403.

Bolwell GP, Wojtaszek P, 1997. Mechanisms for the generation of reactive oxygen species in plant defence – a broad perspective. *Physiological and Molecular Plant Pathology* 51, 347-66.

Choi H-I, Hong J-H, Ha J-O, Kang J-Y, Kim SY, 2000. ABFs, a Family of ABA-responsive Element Binding Factors. *Journal of Biological Chemistry* 275, 1723-30.

Chow BY, Helfer A, Nusinow DA, Kay SA, 2012. ELF3 recruitment to the PRR9 promoter requires other Evening Complex members in the Arabidopsis circadian clock. *Plant Signal Behav* 7, 170-3.

Christianson JA, Wilson IW, Llewellyn DJ, Dennis ES, 2009. The low-oxygen-induced NAC domain transcription factor ANAC102 affects viability of Arabidopsis seeds following low-oxygen treatment. *Plant Physiol* 149(4), 1724-38.

Clough SJ, Bent AF, 1998. Floral dip: a simplified method for Agrobacterium-mediated tranformation of Arabidopsis thaliana. *The Plant Journal* 16(6), 735-743.

Colón-Carmona A, Chen DL, Yeh K-C, Abel S, 2000. Aux/IAA Proteins Are Phosphorylated by Phytochrome in Vitro. *Plant Physiology* 124, 1728-38.

Covington MF, Satchidananda P, Liu XL, Strayer, CA, Wagner DR, Kay SA, 2001. ELF3 modulates resetting of the circadian clock in Arabidopsis. *Plant Cell* 13(6), 1305-1316.

Covington MF, Harmer SL, 2007. The circadian clock regulates auxin signaling and responses in Arabidopsis. *PLoS Biol* 5, e222.

Covington MF, Maloof JN, Straume M, Kay SA, Harmer SL, 2008. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biol* 9, R130.

De Folter S, Immink RGH, Kieffer M, *et al.*, 2005. Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. *The Plant Cell Online* 17, 1424-33.

Dixon LE, Knox K, Kozma-Bognar L, Southern MM, Pokhilko A, Millar AJ, 2011. Temporal repression of core circadian genes is mediated through EARLY FLOWERING 3 in Arabidopsis. *Current Biology* 21, 120-5.

Doyle MR, Davis SJ, Bastow RM, *et al.*, 2002. The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana. *Nature* 419, 74-7.

Dreher KA, Brown J, Saw RE, Callis J, 2006. The Arabidopsis Aux/IAA Protein Family Has Diversified in Degradation and Auxin Responsiveness. *The Plant Cell Online* 18, 699-714.

Edwards KD, Millar A, 2007. Circadian gene expression under different light treatments by transcriptome profiling on microarrays: NASCARRAYS-196. Data set/Database.

Elliott JA, Stetson MH, Menaker M, 1972. Regulation of testis function in golden hamsters: a circadian clock measures photoperiodic time. *Science* 178, 771-3.

Elliott RC, Betzner AS, Huttner E, *et al.*, 1996. AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. *The Plant Cell Online* 8, 155-68.

Elowitz MB, Leibler S, 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335-8.

Eulgem T, Rushton PJ, Robatzek S, Somssich IE, 2000. The WRKY superfamily of plant transcription factors. *Trends in Plant Science* 5, 199-206.

Farré EM, Harmer SL, Harmon FG, Yanovsky MJ, Kay SA, 2005. Overlapping and Distinct Roles of PRR7 and PRR9 in the Arabidopsis Circadian Clock. *Current Biology* 15, 47-54.

Finkelstein R, Gampala SS, Lynch TJ, Thomas TL, Rock CD, 2005. Redundant and distinct functions of the ABA response loci ABA-INSENSITIVE(ABI)5 and ABRE-BINDING FACTOR (ABF)3. *Plant Molecular Biology* 59, 253-67.

Gendrel AV, Lippman Z, Yordan C, Colot V, Martienssen R.A, 2002. Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene DDM1. *Science* 297, 1871-1873.

Gendron JM, Pruneda-Paz JL, Doherty CJ, Gross AM, Kang SE, Kay SA, 2012. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proc Natl Acad Sci U S A* 109, 3167-72.

Gietz RD, Schiestl RH, 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2, 31-4.

Goda H, Sasaki E, Akiyama K, *et al.*, 2008. The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J* 55(3), 526-42.

Gong W, Shen Y-P, Ma L-G, *et al.*, 2004. Genome-Wide ORFeome Cloning and Analysis of Arabidopsis Transcription Factor Genes. *Plant Physiology* 135, 773-82.

Gould PD, Locke JC, Larue C, *et al.*, 2006. The molecular basis of temperature compensation in the Arabidopsis circadian clock. *Plant Cell* 18, 1177-87.

Green RM, Tobin EM, 1999. Loss of the circadian clock-associated protein 1 in Arabidopsis results in altered clock-regulated gene expression. *Proc Natl Acad Sci U S A* 96, 4176-9.

Grill E, Himmelbach A, 1998. ABA signal transduction. *Curr Opin Plant Biol* 1, 412-8.

Gutierrez L, Van Wuytswinkel O, Castelain M, Bellini C, 2007. Combined networks regulating seed maturation. *Trends Plant Sci* 12, 294-300.

Haring M, Offermann S, Danker T, Horst I, Peterhansel C, Stam M, 2007. Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods* 3:11.

Hazen SP, Schultz TF, Pruneda-Paz JL, Borevitz JO, Ecker JR, Kay SA, 2005. LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. *Proc Natl Acad Sci U S A* 102, 10387-92.

Helfer A, Nusinow DA, Chow BY, Gehrke AR, Bulyk ML, Kay SA, 2011. LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the Arabidopsis core clock. *Curr Biol* 21, 126-33.

Helliwell CA, Wood CC, Robertson M, James Peacock W, Dennis ES, 2006. The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *Plant J* 46, 183-92.

Hicks KA, Albertson TM, Wagner DR, 2001. EARLY FLOWERING3 encodes a novel protein that regulates circadian clock function and flowering in Arabidopsis. *Plant Cell* 13, 1281-92.

Hruz T, Laule O, Szabo G, *et al.*, 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* 2008, Epub Article ID 420747.

Hu Y, Xie Q, Chua N-H, 2003. The Arabidopsis Auxin-Inducible Gene ARGOS Controls Lateral Organ Size. *The Plant Cell Online* 15, 1951-61.

Huang W, Perez-Garcia P, Pokhilko A, *et al.*, 2012. Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. *Science* 336, 75-9.

Huang W, Ramsey KM, Marcheva B, Bass J, 2011. Circadian rhythms, sleep, and metabolism. *The Journal of Clinical Investigation* 121, 2133-41.

Hudson ME, Quail PH, 2003. Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. *Plant Physiol* 133, 1605-16.

Iglesias MJ, Terrile MC, Bartoli CG, D'ippolito S, Casalongue CA, 2010. Auxin signaling participates in the adaptative response against oxidative stress and salinity by interacting with redox metabolism in Arabidopsis. *Plant Molecular Biology* 74, 215-22.

Ito S, Song YH, Josephson-Day AR, *et al.*, 2012. FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. *Proceedings of the National Academy of Sciences* 109, 3582-7.

Jakoby M, Weisshaar B, Dröge-Laser W, *et al.*, 2002. bZIP transcription factors in Arabidopsis. *Trends in Plant Science* 7, 106-11.

James AB, Syed NH, Bordage S, *et al.*, 2012. Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant Cell* 24, 961-81.

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, Mcginnis S, Madden TL, 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res* 36, W5-W9.

Jung JH, Park CM, 2011. Auxin modulation of salt stress signaling in Arabidopsis seed germination. *Plant Signal Behav* 6, 1198-200.

Kang J-Y, Choi H-I, Im M-Y, Kim SY, 2002. Arabidopsis Basic Leucine Zipper Proteins That Mediate Stress-Responsive Abscisic Acid Signaling. *The Plant Cell Online* 14, 343-57.

Khanna R, Kikis EA, Quail PH, 2003. EARLY FLOWERING 4 functions in phytochrome B-regulated seedling de-etiolation. *Plant Physiol* 133, 1530-8.

Kidokoro S, Maruyama K, Nakashima K, *et al.*, 2009. The Phytochrome-Interacting Factor PIF7 Negatively Regulates DREB1 Expression under Circadian Control in Arabidopsis. *Plant Physiology* 151, 2046-57.

Kilian J, Whitehead D, Horak J, Wanke D *et al.*, 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* 50(2), 347-63.

Kim J, Harter K, Theologis A, 1997. Protein–protein interactions among the Aux/IAA proteins. *Proceedings of the National Academy of Sciences* 94, 11786-91.

Kim SY, Ma J, Perret P, Li Z, Thomas TL, 2002. Arabidopsis ABI5 subfamily members have distinct DNA-binding and transcriptional activities. *Plant Physiol* 130, 688-97.

Kim WY, Fujiwara S, Suh SS, *et al.*, 2007. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature* 449, 356-60.

Kleine T, Kindgren P, Benedict C, Hendrickson L *et al*., 2007. Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of Arabidopsis to high irradiance. *Plant Physiol* 144(3), 1391-406.

Krizek B, 2009. AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate Arabidopsis floral growth and patterning. *Plant Physiol* 150, 1916-29.

Krizek BA, 2011. Auxin regulation of Arabidopsis flower development involves members of the AINTEGUMENTA-LIKE/PLETHORA (AIL/PLT) family. *J Exp Bot* 62, 3311-9.

Krizek BA, Prost V, Macias A, 2000. AINTEGUMENTA promotes petal identity and acts as a negative regulator of AGAMOUS. *Plant Cell* 12, 1357-66.

Kurup S, Jones HD, Holdsworth MJ, 2000. Interactions of the developmental regulator ABI3 with proteins identified from developing Arabidopsis seeds. *Plant J* 21, 143-55.

Lee C-M, Thomashow MF, 2012. Photoperiodic regulation of the C-repeat binding factor (CBF) cold acclimation pathway and freezing tolerance in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 109, 15054-9.

Lee J, Lee I, 2010. Regulation and function of SOC1, a flowering pathway integrator. *J Exp Bot* 61, 2247-54.

Leivar P, Monte E, Al-Sady B, *et al.*, 2008. The Arabidopsis phytochrome-interacting factor PIF7, together with PIF3 and PIF4, regulates responses to prolonged red light by modulating phyB levels. *Plant Cell* 20, 337-52.

Levy Y, Hanan E, Solomon B, Becker OM, 2001. Helix-coil transition of PrP106-126: molecular dynamic study. *Proteins* 45, 382-96.

Li J, Krichevsky A, Vaidya M, Tzfira T, Citovsky V, 2005. Uncoupling of the functions of the Arabidopsis VIP1 protein in transient and stable plant genetic transformation by Agrobacterium. *Proceedings of the National Academy of Sciences of the United States of America* 102, 5733-8.

Li L, Ljung K, Breton G, *et al.*, 2012a. Linking photoreceptor excitation to changes in plant architecture. *Genes Dev* 26, 785-90.

Li MD, Li CM, Wang Z, 2012b. The role of circadian clocks in metabolic disease. *Yale J Biol Med* 85, 387-401.

Liu XL, Covington MF, Fankhauser C, Chory J, Wagner DR, 2001. ELF3 encodes a circadian clock-regulated nuclear protein that functions in an Arabidopsis PHYB signal transduction pathway. *Plant Cell* 13, 1293-304.

Locke JCW, Kozma-Bognar L, Gould PD, Fehér B, Kevei E, Nagy F, Turner MS, Hall A, Millar AJ, 2006. Experimental validation of a predicted feedback loop in the multi-oscillator clock of Arabidopsis thaliana. *Mol Syst Biol* 2, 59.

Lu SX, Webb CJ, Knowles SM, Kim SH, Wang Z, Tobin EM, 2012. CCA1 and ELF3 interact in the control of hypocotyl length and flowering time in Arabidopsis. *Plant Physiol* 158, 1079-88.

Lu W, Meng Q-J, Tyler NJC, Stokkan K-A, Loudon ASI, 2010. A Circadian Clock Is Not Required in an Arctic Mammal. *Current Biology* 20, 533-7.

Makino S, Matsushika A, Kojima M, Oda Y, Mizuno T, 2001. Light response of the circadian waves of the APRR1/TOC1 quintet: when does the quintet start singing rhythmically in Arabidopsis? *Plant Cell Physiol* 42, 334-9.

Makino S, Matsushika A, Kojima M, Yamashino T, Mizuno T, 2002. The APRR1/TOC1 quintet implicated in circadian rhythms of Arabidopsis thaliana: I. Characterization with APRR1-overexpressing plants. *Plant Cell Physiol* 43, 58-69.

Mallappa C, Singh A, Ram H, Chattopadhyay S, 2008. GBF1, a transcription factor of blue light signaling in Arabidopsis, is degraded in the dark by a proteasome-mediated pathway independent of COP1 and SPA1. *J Biol Chem* 283, 35772-82.

Mallappa C, Yadav V, Negi P, Chattopadhyay S, 2006. A basic leucine zipper transcription factor, G-box-binding factor 1, regulates blue light-mediated photomorphogenic growth in Arabidopsis. *J Biol Chem* 281, 22190-9.

Martinez-Garcia JF, Huq E, Quail PH, 2000. Direct targeting of light signals to a promoter element-bound transcription factor. *Science* 288, 859-63.

Más P, Alabadí D, Yanovsky MJ, Oyama T, Kay SA, 2003a. Dual Role of TOC1 in the Control of Circadian and Photomorphogenic Responses in Arabidopsis. *The Plant Cell Online* 15, 223-36.

Más P, Kim WY, Somers DE, Kay SA, 2003b. Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana. *Nature* 426, 567-70.

Matsushika A, Makino S, Kojima M, Mizuno T, 2000. Circadian Waves of Expression of the APRR1/TOC1 Family of Pseudo-Response Regulators in Arabidopsis thaliana: Insight into the Plant Circadian Clock. *Plant and Cell Physiology* 41, 1002-12.

McClung CR, 2006. Plant circadian rhythms. *Plant Cell* 18, 792-803.

McClung CR, Hsu M, Painter JE, Gagne JM, Karlsberg SD, Salome PA, 2000. Integrated temporal regulation of the photorespiratory pathway. Circadian regulation of two Arabidopsis genes encoding serine hydroxymethyltransferase. *Plant Physiol* 123, 381-92.

McWatters HG, Kolmos E, Hall A, *et al.*, 2007. ELF4 is required for oscillatory properties of the circadian clock. *Plant Physiol* 144, 391-401.

Meister RJ, Williams LA, Monfared MM, *et al.*, 2004. Definition and interactions of a positive regulatory element of the Arabidopsis INNER NO OUTER promoter. *Plant J* 37, 426-38.

Merlot S, Gosti F, Guerrier D, Vavasseur A, Giraudat J, 2001. The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. *The Plant Journal* 25, 295-303.

Michael TP, Mockler TC, Breton G, *et al.*, 2008. Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet* 4, e14.

Michel K, Abderhalden O, Bruggmann R, Dudler R, 2006. Transcriptional changes in powdery mildew infected wheat and Arabidopsis leaves undergoing syringolin-triggered hypersensitive cell death at infection sites. *Plant Mol Biol* 62, 561-578.

Millar AJ, Carré IA, Strayer CA, Chua NH, Kay SA, 1995. Circadian clock mutants in Arabidopsis identified by luciferase imaging. *Science* 267, 1161-3.

Mizoguchi T, Wheatley K, Hanzawa Y, *et al.*, 2002. LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. *Dev Cell* 2, 629-41.

Mizukami Y, Fischer RL, 2000. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A* 97, 942-7.

Mockler TC, Michael TP, Priest HD, *et al.*, 2007. The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harb Symp Quant Biol* 72, 353-63.

Monfared MM, Simon MK, Meister RJ, *et al.*, 2011. Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *The Plant Journal* 66, 1020-31.

Mönke G, Altschmied L, Tewes A, *et al.*, 2004. Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* 219, 158-66.

Mönke G, Seifert M, Keilwagen J, *et al.*, 2012. Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon. *Nucleic Acids Res* 40, 8240-54.

Monteiro PT, Mendes ND, Teixeira MC, *et al.*, 2008. YEASTRACT-DISCOVERER: new tools to improve the analysis of transcriptional regulatory associations in Saccharomyces cerevisiae. *Nucleic Acids Res* 36, D132-D6.

Moubayidin L, Di Mambro R, Sabatini S, 2009. Cytokinin-auxin crosstalk. *Trends Plant Sci* 14, 557-62.

Nakamichi N, Kiba T, Henriques R, Mizuno T, Chua NH, Sakakibara H, 2010. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *Plant Cell* 22, 594-605.

Nakamichi N, Kita M, Ito S, Yamashino T, Mizuno T, 2005. PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of Arabidopsis thaliana. *Plant Cell Physiol* 46, 686-98.

Nakamura S, Lynch TJ, Finkelstein RR, 2001. Physical interactions between ABA response loci of Arabidopsis. *Plant J* 26, 627-35.

Nelson MS, Xie YH, Asrani AC, Gupta P, 2006. A reliable strategy for validation of real-time quantitative RT-PCR amplicons from linearly amplified human hematopoietic stem cell mRNA. *BLOOD* 108, 131B.

Ni M, Tepperman JM, Quail PH, 1998. PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell* 95, 657-67.

Noguchi T, Fujioka S, Takatsuto S, *et al.*, 1999a. Arabidopsis det2 is defective in the conversion of (24R)-24-methylcholest-4-En-3-one to (24R)-24-methyl-5alpha-cholestan-3-one in brassinosteroid biosynthesis. *Plant Physiol* 120, 833-40.

Nole-Wilson S, Krizek BA, 2000. DNA binding properties of the Arabidopsis floral development protein AINTEGUMENTA. *Nucleic Acids Res* 28, 4076-82.

Nusinow DA, Helfer A, Hamilton EE, *et al.*, 2011. The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* 475, 398-402.

Obayashi T, Nishida K, Kasahara K, Kinoshita K, 2011. ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant Cell Physiol* 52, 213-9.

Oh E, Kang H, Yamaguchi S, *et al.*, 2009. Genome-Wide Analysis of Genes Targeted by PHYTOCHROME INTERACTING FACTOR 3-LIKE5 during Seed Germination in Arabidopsis. *The Plant Cell Online* 21, 403-19.

Olsen AN, Ernst HA, Leggio LL, Skriver K, 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* 10, 79-87.

Onai K, Ishiura M, 2005. PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. *Genes Cells* 10, 963-72.

Panda S, Antoch MP, Miller BH, *et al.*, 2002. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* 109, 307-20.

Park J, Lee N, Kim W, Lim S, Choi G, 2011. ABI3 and PIL5 Collaboratively Activate the Expression of SOMNUS by Directly Binding to Its Promoter in Imbibed Arabidopsis Seeds. *The Plant Cell Online* 23, 1404-15.

Pauwels L, Morreel K, De Witte E, Lammertyn F, Van Montagu M, Boerjan W, Inzé D, Goossens A, 2008. *Proceedings of the National Academy of Sciences of the United States of America* 105(4), 1380-1385.

Paz-Ares J, 2002. REGIA, an EU project on functional genomics of transcription factors from Arabidopsis thaliana. *Comp Funct Genomics* 3, 102-8.

Penfield S, Hall A, 2009. A Role for Multiple Circadian Clock Genes in the Response to Signals That Break Seed Dormancy in Arabidopsis. *The Plant Cell Online* 21, 1722-32.

Peng X, Hu Y, Tang X, *et al.*, 2012. Constitutive expression of rice WRKY30 gene increases the endogenous jasmonic acid accumulation, PR gene expression and resistance to fungal pathogens in rice. *Planta* 236, 1485-98.

Petersen M, Brodersen P, Naested H, *et al.*, 2000. Arabidopsis MAP Kinase 4 Negatively Regulates Systemic Acquired Resistance. *Cell* 103, 1111-20.

Picot E, Krusche P, Tiskin A, Carré I, Ott S, 2010. Evolutionary analysis of regulatory sequences (EARS) in plants. *Plant J* 64, 165-76.

Pokhilko A, Fernández AP, Edwards KD, Southern MM, Halliday KJ, Millar AJ, 2012. The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Mol Syst Biol* 8, 574.

Pruneda-Paz JL, Breton G, Para A, Kay SA, 2009. A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. *Science* 323, 1481-5.

Reed JW, 2001. Roles and activities of Aux/IAA proteins in Arabidopsis. *Trends Plant Sci* 6, 420-5.

Sambrook J, Fritsch EF, Maniatis T, 1989. *Molecular Cloning: A Laboratory Manual*. Volume 1. Cold Spring Harbor, NY, USA.

Scarpeci T, Zanor M, Carrillo N, Mueller-Roeber B, Valle E, 2008. Generation of superoxide anion in chloroplasts of Arabidopsis thaliana during active photosynthesis: a focus on rapidly induced genes. *Plant Molecular Biology* 66, 361-78.

Schaffer R, Ramsay N, Samach A, *et al.*, 1998. The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93, 1219-29.

Schellenberg B, Ramel C, Dudler R, 2010. Pseudomonas syringae virulence factor syringolin A counteracts stomatal immunity by proteasome inhibition. *Mol Plant Microbe Interact* 23, 1287-93.

Schindler U, Terzaghi W, Beckmann H, Kadesch T, Cashmore AR, 1992. DNA binding site preferences and transcriptional activation properties of the Arabidopsis transcription factor GBF1. *EMBO J* 11, 1275-89.

Scholl RL, May ST, Ware DH, 2000. Seed and Molecular Resources for Arabidopsis. *Plant Physiology* 124, 1477-80.

Sehgal A, Joiner W, Crocker A, *et al.*, 2007. Molecular Analysis of Sleep: Wake Cycles in Drosophila. *Cold Spring Harb Symp Quant Biol* 72, 557-64.

Shibasaki K, Uemura M, Tsurumi S, Rahman A, 2009. Auxin response in Arabidopsis under cold stress: underlying molecular mechanisms. *Plant Cell* 21, 3823-38.

Somers DE, Webb AA, Pearson M, Kay SA, 1998. The short-period mutant, toc1-1, alters circadian clock regulation of multiple outputs throughout development in Arabidopsis thaliana. *Development* 125, 485-94.

Spensley M, Kim JY, Picot E, *et al.*, 2009. Evolutionarily conserved regulatory motifs in the promoter of the Arabidopsis clock gene LATE ELONGATED HYPOCOTYL. *Plant Cell* 21, 2606-23.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M, 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 34, D535-9.

Strayer C, Oyama T, Schultz TF, *et al.*, 2000. Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science* 289, 768-71.

Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G, 2001. CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410, 1116-20.

Teixeira MC, Monteiro P, Jain P, *et al.*, 2006. The YEASTRACT database: a tool for the analysis of transcription regulatory associations in Saccharomyces cerevisiae. *Nucleic Acids Res* 34, D446-D51.

Truman WM, Bennett MH, Turnbull CG, Grant MR, 2010. Arabidopsis auxin mutants are compromised in systemic acquired resistance and exhibit aberrant accumulation of various indolic compounds. *Plant Physiol* 152, 1562-73.

Tsugama D, Liu S, Takano T, 2012. A bZIP protein, VIP1, is a regulator of osmosensory signaling in Arabidopsis. *Plant Physiology*.

Vanneste S, De Rybel B, Beemster GT, *et al.*, 2005. Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in Arabidopsis thaliana. *Plant Cell* 17(11), 3035-50.

Wang KL-C, Li H, Ecker JR, 2002. Ethylene Biosynthesis and Signaling Networks. *The Plant Cell Online* 14, S131-S51.

Wang W, Malcolm BA, 1999. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *Biotechniques* 26, 680-2.

Wang XE, Basnayake BMVS, Zhang H, *et al.*, 2009. The Arabidopsis ATAF1, a NAC Transcription Factor, Is a Negative Regulator of Defense Responses Against Necrotrophic Fungal and Bacterial Pathogens. *Molecular Plant-Microbe Interactions* 22, 1227-38.

Wang YH, Irving HR, 2011. Developing a model of plant hormone interactions. *Plant Signal Behav* 6, 494-500.

Wang ZY, Kenigsbuch D, Sun L, Harel E, Ong MS, Tobin EM, 1997. A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* 9, 491-507.

Wang ZY, Tobin EM, 1998. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93, 1207-17.

Zagotta MT, Hicks KA, Jacobs CI, Young JC, Hangarter RP, Meeks-Wagner DR, 1996. The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *Plant J* 10, 691-702.

Zhang X, Henriques R, Lin SS, Niu QW, Chua NH, 2006. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nat Protoc* 2, 641-6.

Zhou X-Y, Song L, Xue H-W, 2012. Brassinosteroids Regulate the Differential Growth of Arabidopsis Hypocotyls through Auxin Signaling Components IAA19 and ARF7. *Molecular Plant* 6(3), 887-904.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, Felix G, 2006. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell* 125(4), 749-60.